Konzeption und Implementierung von Gleichheit für einen tableau–basierten Theorem–Beweiser

Bernhard Beckert

Universität Karlsruhe
Institut für Logik, Komplexität und Deduktionssysteme

Kurzzusammenfassung

Tableau–basierte Theorembeweiser lassen sich an viele der nicht–klassischen Logiken anpassen, die zur Zeit im Bereich der KI untersucht werden. Gleichheit ist sowohl für die klassische als auch für die nicht–klassische Prädikatenlogik von großer Bedeutung, da sie die Ausdrucksstärke der Objektsprache erhöht. Leider sind alle bisherigen Versuche, den Tableaukalkül um die Behandlung von Gleichheit zu erweitern, mehr oder weniger experimenteller Natur, und es zeigt sich, daß sie für die praktische Anwendung nicht geeignet sind. In der vorliegenden Arbeit wird ein weiterführender Ansatz vorgestellt, der bessere Ergebnisse liefert und die Voraussetzungen für weitere Entwicklungen schafft. Die Probleme, die sich inbesondere daraus ergeben, daß der Tableaukalkül als Rahmen gewählt wurde, werden aufgezeigt. Korrektheit und Vollständigkeit des neuen Ansatzes werden bewiesen. Darüberhinaus wird seine Implementierung als Teil des Tableaubeweisers $_3T^AP$ beschrieben.

Inhaltsverzeichnis

Kapitel 1

Einführung

Der Tableaukalkül hat als logische Basis für einen automatischen Theorembeweiser drei wesentliche Vorteile: erstens gibt es keine Einschränkung auf bestimmte Normalformen, zweitens kann der Tableaukalkül an viele nicht–klassische Logiken angepaßt werden, und drittens können für Formeln, die keine Tautologien sind, Gegenbeispiele angegeben werden.

Gleichheit ist sowohl für die klassische als auch für die nicht–klassische Prädikatenlogik von großer Bedeutung, da sie die Ausdrucksstärke der Objektsprache erhöht. Leider sind alle bisherigen Versuche, den Tableaukalkül um die Behandlung von Gleichheit zu erweitern, mehr oder weniger experimenteller Natur, und es zeigt sich, daß sie für die praktische Anwendung nicht geeignet sind. In der vorliegenden Arbeit wird ein weiterführender Ansatz vorgestellt, der bessere Ergebnisse liefert und die Voraussetzungen für weitere Entwicklungen schafft. Die Probleme, die sich inbesondere daraus ergeben, daß der Tableaukalkül als Rahmen gewählt wurde, werden aufgezeigt. Korrektheit und Vollständigkeit des neuen Ansatzes werden bewiesen. Darüberhinaus wird seine Implementierung als Teil des Tableaubeweisers $_{3}T^{A}P$ beschrieben.

Der Beweiser $_{3}T^{A}P$ wird seit 1989 von Reiner Hähnle am Institut für Logik, Komplexität und Deduktionssysteme der Universität Karlsruhe entwickelt. Das System ist in [Hähnle, 1990a], [Hähnle, 1990b] und [Hähnle, 1991] beschrieben.

Die Entwicklung von $_{3}T^{A}P$ und diese Arbeit wurden von IBM Deutschland gefördert, im Rahmen eines gemeinsamen Projektes (TCG Projekt) der Universität Karlsruhe und des Institutes für Wissensbasierte Systeme von IBM Deutschland in Heidelberg.

Dieser Arbeit lag die zweiwertige Version von $_{3}T^{A}P$ zugrunde. Die Möglichkeit, mehrwertige Logiken zu behandeln, — eine wesentliche Eigenschaft von $_{3}T^{A}P$ — wurde hier nicht berücksichtigt.

In Kapitel 2 wird zunächst eine kurze Einführung in die Theorie der Gleichheit im Tableaukalkül gegeben, und die aus der Literatur bekannten Ansätze werden beschrieben. Die Konzeption des schließlich implementierten Verfahrens wird in den Kapiteln 3 und 4 entwickelt. Vor– und Nachteile des Verfahrens werden betrachtet und seine Korrektheit und Vollständigkeit bewiesen. In Kapitel 5 ist die Implementierung des Verfahrens beschrieben. Näher eingegangen wird dabei auf die verwendeten Datenstrukturen, die Struktur des in $_{3}T^{A}P$ integrierten Gleichheits–Moduls, einige besonders interessante und wichtige Algorithmen und die Form, in der Beweise ausgegeben werden. Schließlich wird in Kapitel 6 ein kurzes Fazit gezogen, und mögliche Verbesserungen des vorgestellten Ansatzes und Erweiterungen der Implementierung werden vorgeschlagen. Der Anhang enthält einige Beispiele für Tableaubeweise, die mit dem um Gleichheit erweiterten $_{3}T^{A}P$–System generiert wurden.

Kapitel 2

Theorie der Gleichheit im Tableau–Kalkül

2.1 Prädikatenlogik mit freien Variablen und Gleichheit

2.1.1 Syntax

Eine ausführliche Darstellung der Syntax der Prädikatenlogik, insbesondere auch der mit freien Variablen und Gleichheit, findet sich in [Fitting, 1990]. Dort sind die Menge der Terme **Term**, die Menge der Variablen **Var** und die Menge der Formeln **Form** für die verschiedenen Versionen der Prädikatenlogik (in der üblichen Weise) definiert — mit und ohne freien Variablen, mit und ohne Gleichheit. Die Gleichheit wird auf syntaktischer Ebene mit dem Prädikatensymbol $\approx$ bezeichnet.

Zum Aufbau eines Tableaus werden die folgenden Erweiterungsregeln verwendet. Sie sind nach den Schemata gegliedert, denen sie folgen. Zunächst die α– und β–Regeln:

Definition 2.1.1 **(α– und β–Regeln)**

Schemata für die α– und β–Regeln:

$$\frac{\alpha}{\begin{array}{c}\alpha_1\\ \alpha_2\end{array}} \qquad\qquad \frac{\beta}{\beta_1 \mid \beta_2}$$

Die α–Regeln:

$$\frac{\mathsf{T}\,(X_1 \wedge X_2)}{\begin{array}{c}\mathsf{T}\,X_1\\ \mathsf{T}\,X_2\end{array}} \qquad \frac{\mathsf{F}\,(X_1 \vee X_2)}{\begin{array}{c}\mathsf{F}\,X_1\\ \mathsf{F}\,X_2\end{array}} \qquad \frac{\mathsf{F}\,(X_1 \supset X_2)}{\begin{array}{c}\mathsf{T}\,X_1\\ \mathsf{F}\,X_2\end{array}}$$

$$\frac{\mathsf{T}\,\neg X}{\mathsf{F}\,X} \qquad\qquad \frac{\mathsf{F}\,\neg X}{\mathsf{T}\,X}$$

Die β–Regeln:

$$\frac{\mathsf{F}\,(X_1 \wedge X_2)}{\mathsf{F}\,X_1 \mid \mathsf{F}\,X_2} \qquad \frac{\mathsf{T}\,(X_1 \vee X_2)}{\mathsf{T}\,X_1 \mid \mathsf{T}\,X_2} \qquad \frac{\mathsf{T}\,(X_1 \supset X_2)}{\mathsf{F}\,X_1 \mid \mathsf{T}\,X_2}$$

Es gibt zwei Versionen der γ– und der δ–Regeln, eine für das Tableau mit und eine für das Tableau ohne freie Variablen.

Definition 2.1.2 **(γ– und δ–Regeln ohne freie Variablen)**

Schemata für die γ– und die δ–Regeln:

$$\frac{\gamma}{\gamma(t)} \qquad\qquad \frac{\delta}{\delta(c)}$$

Die γ–Regeln:

$$\frac{\mathsf{T}\,(\forall x)Z(x)}{\mathsf{T}\,Z(t)} \qquad\qquad \frac{\mathsf{F}\,(\exists x)Z(x)}{\mathsf{F}\,Z(t)}$$

Die δ–Regeln:

$$\frac{\mathsf{F}\,(\forall x)Z(x)}{\mathsf{F}\,Z(c)} \qquad\qquad \frac{\mathsf{T}\,(\exists x)Z(x)}{\mathsf{T}\,Z(c)}$$

Dabei bezeichnet t einen beliebigen geschlossenen Term und c eine neue Skolemkonstante.

Definition 2.1.3 (γ– und δ–Regeln mit freien Variablen)

Schemata für die γ– und die δ–Regeln:

$$\frac{\gamma}{\gamma(y)} \qquad\qquad \frac{\delta}{\delta(f(x_1,\ldots,x_n))}$$

Die γ–Regeln:

$$\frac{\mathsf{T}\,(\forall x)Z(x)}{\mathsf{T}\,Z(y)} \qquad\qquad \frac{\mathsf{F}\,(\exists x)Z(x)}{\mathsf{F}\,Z(y)}$$

Die δ–Regeln:

$$\frac{\mathsf{F}\,(\forall x)Z(x)}{\mathsf{F}\,Z(f(x_1,\ldots,x_n))} \qquad\qquad \frac{\mathsf{T}\,(\exists x)Z(x)}{\mathsf{T}\,Z(f(x_1,\ldots,x_n))}$$

Dabei bezeichnet y eine beliebige freie Variable, f eine neue Skolemfunktion, und $x_1,\ldots,x_n$ sind die in δ frei auftretenden Variablen.

Die δ–Regel für das Tableau mit freien Variablen weicht von der in [Fitting, 1990] gegebenen ab. Diese neue, beim Beweiser $_3T^AP$ verwendete Regel stellt eine Vereinfachung dar, da nicht alle freien Variablen des Tableaus zu Argumenten der Skolemfunktion werden. Die Korrektheit dieser Regel ist in [Hähnle & Schmitt, 1991] bewiesen.

Die in den Definitionen 2.1.1, 2.1.2 und 2.1.3 eingeführten Regeln beziehen sich auf signierte Formeln. In der Literatur wird dagegen zumeist ein Tableaukalkül mit nicht signierten Formeln eingeführt, was allerdings keinen Einfluß auf die weiteren Betrachtungen hat. Alle Ergebnisse gelten gleichermaßen für den Tableaukalkül mit Vorzeichen wie für den ohne.

Folgendes ist allerdings zu beachten: des öfteren werden über signierte Formeln Aussagen gemacht werden, die nur für nicht signierte Formeln einen Sinn ergeben — so beispielsweise Aussagen über die Erfüllbarkeit, die nur für nicht signierte Formeln definiert wird. In solchen Fällen ist eine signierte Formel durch die mit ihr korrespondierende nicht signierte Formel zu ersetzen:

Definition 2.1.4 (Mit einer signierten Formel korrespondierende nicht signierte Formel)

Die zu einer signierten Formel korrespondierende nicht signierte Formel erhält man in folgender Weise:

- *Eine positive Signatur wird entfernt:* $\mathsf{T}\,X \rightsquigarrow X$.
- *Eine negative Signatur wird durch $\neg$ ersetzt:* $\mathsf{F}\,X \rightsquigarrow \neg X$.

2.1.2 Substitution der freien Variablen

Da Tableaus mit freien Variablen betrachtet werden, muß vor der Definition ihrer Erfüllbarkeit der Begriff der Substitution eingeführt werden. Er ist für das weitere Vorgehen von zentraler Bedeutung. Daher ist ihm ein eigener Abschnitt gewidmet.

Definition 2.1.5 (Substitution)

Eine Substitution ist eine Abbildung $\sigma : \mathbf{Var} \longrightarrow \mathbf{Term}$ der freien Variablen auf die Terme.

Ist die Menge der Variablen $\{x_1,\ldots,x_n\}$, die durch eine Substitution σ nicht auf sich abgebildet werden, endlich, so kann σ durch

$$\sigma = \{x_1 \leftarrow t_1,\ldots,x_n \leftarrow t_n\}$$

dargestellt werden. Dabei bezeichnet t_i $(1 \le i \le n)$ das Bild der Variablen x_i unter σ.

Definition 2.1.6 (Anwendung einer Substitution)

Die Anwendung $t\sigma$ einer Substitution σ auf einen Term t erhält man, indem man alle in t frei auftretenden Variablen x durch $\sigma(x)$ ersetzt.

Alle Ersetzungen sind simultan auszuführen.

Beispiel 2.1.7

Sei $\sigma = \{x \leftarrow f(x,y),\ y \leftarrow h(a),\ z \leftarrow a\}$. Dann ist

$$h(g(x),y)\sigma = h(g(f(x,y)),h(a)).$$

Definition 2.1.8 (Grundsubstitution)

Eine Substitution σ heißt Grundsubstitution, wenn

1. *durch σ alle freien Variablen substituiert werden, und*
2. *die Terme $t_1, \ldots, t_n$ keine freien Variablen enthalten.*

In den weiteren Kapiteln werden auch solche Substitutionen als Grundsubstitution aufgefaßt, die alle in einem betrachteten Tableau auftretenden Variablen belegen.

Definition 2.1.9 (Kanonische Substitution)

Eine Substitution

$$\sigma = \{x_1 \leftarrow t_1, \ldots, x_n \leftarrow t_n\},$$

heißt kanonisch, wenn keine der Variablen x_i in einem der Terme t_j vorkommt ($1 \leq i, j \leq n$).

Bei der Anwendung einer kanonische Substitution müssen die Einzelsubstitutionen nicht parallel sondern können auch nacheinander ausgeführt werden.

Beispiel 2.1.10

Kanonisch sind beispielsweise die Substitutionen

- $\{x \leftarrow y,\ z \leftarrow y\}$
- $\{x \leftarrow a,\ y \leftarrow f(b)\}$.

Nicht kanonisch sind

- $\{x \leftarrow y,\ y \leftarrow a\}$.
- $\{x \leftarrow f(x)\}$.

Definition 2.1.11 (Allgemeinere Substitution, Spezialisierung)

Eine Substitution σ heißt allgemeiner als eine Substitution τ, wenn es eine Substitution σ' gibt mit

$$\tau = \sigma' \circ \sigma.$$

In diesem Falle heißt τ eine Spezialisierung von σ.

In Zeichen: $\sigma \succ \tau$.

Beispiel 2.1.12

$\sigma = \{x \leftarrow f(y)\}$ ist allgemeiner als $\{x \leftarrow f(a),\ y \leftarrow a\}$. σ ist jedoch nicht allgemeiner als $\{x \leftarrow f(a)\}$.

		Vereinigung
$\{x \leftarrow a\}$	$\{y \leftarrow b\}$	$\{x \leftarrow a, y \leftarrow b\}$
$\{x \leftarrow a\}$	$\{x \leftarrow a\}$	$\{x \leftarrow a\}$
$\{x \leftarrow f(y)\}$	$\{y \leftarrow a\}$	$\{x \leftarrow f(a), y \leftarrow a\}$
$\{x \leftarrow f(y)\}$	$\{x \leftarrow f(a)\}$	$\{x \leftarrow f(a), y \leftarrow a\}$
$\{x \leftarrow a\}$	$\{x \leftarrow b\}$	sind nicht kompatibel
$\{x \leftarrow f(y)\}$	$\{x \leftarrow f(a), y \leftarrow b\}$	sind nicht kompatibel

Tabelle 2.1: Beispiele für die Vereinigung von Substitutionen

Definition 2.1.13 (Kompatibilität, Vereinigung, kanonische Vereinigung)

Zwei Substitutionen τ_1 und τ_2 heißen kompatibel, wenn sie eine gemeinsame Spezialisierung besitzen, wenn also eine Substitution σ existiert, so daß sowohl τ_1 als auch τ_2 allgemeiner ist als σ.

Ist darüberhinaus σ die allgemeinste gemeinsame Spezialisierung von τ_1 und τ_2, so heißt σ eine Vereinigung von τ_1 und τ_2.

Unter einer kanonischen Vereinigung ist eine allgemeinste kanonische gemeinsame Spezialisierung von Substitutionen zu verstehen.

Beispiel 2.1.14

In Tabelle 2.1 sind einige Beispiele für die Vereinigung von Substitutionen angegeben.

2.1.3 Semantik

In diesem Abschnitt sind nur die wichtigsten Definitionen für die Semantik der Prädikatenlogik mit Gleichheit wiedergegeben. Wegen einer ausführlichen Darstellung sei wiederum auf [Fitting, 1990] verwiesen.

Die Modelle der Prädikatenlogik mit Gleichheit, die normalen Modelle, haben im Vergleich zu den Modellen der Prädikatenlogik ohne Gleichheit die Eigenschaft, daß in ihnen das Prädikatensymbol $\approx$ in der gewünschten Weise interpretiert wird:

Definition 2.1.15 (Normales Modell)

Ein Modell $\mathbf{M} = (\mathbf{D}, \mathbf{I})$ heißt normal, wenn die Interpretation des Prädikatensymbols $\approx^{\mathbf{I}}$ die Gleichheitsrelation auf der Grundmenge $\mathbf{D}$ ist.

Definition 2.1.16 (Kanonisches Modell)

Ein normales Modell $\mathbf{M} = (\mathbf{D}, \mathbf{I})$ heißt kanonisch, wenn für jedes $d \in \mathbf{D}$ ein Term t existiert, so daß $t^{\mathbf{I}} = d$ ist.

Wie der folgende Satz zeigt, bilden die kanonische Modelle das Analogon zu den Herbrand–Modellen der Prädikatenlogik ohne Gleichheit.

Satz 2.1.17

Gibt es ein normales Modell, in dem die Menge von Sätzen S erfüllt ist, dann gibt auch ein kanonisches Modell, in dem S erfüllt ist.

Beweis:

Beweis zu Theorem 8.6.3 in [Fitting, 1990]. ■

Aus diesem Satz folgt, daß es genügt, sich bei der Betrachtung der Erfüllbarkeit einer Formel und damit auch der Erfüllbarkeit eines Tableaus auf kanonische Modelle zu beschränken.

Definition 2.1.18 (Erfüllbarkeit eines Tableaus)

Ein Tableau T heißt erfüllbar, wenn es eine kanonische Interpretation $\mathbf{M}$ gibt, so daß es für jede Grundsubstitution σ einen Ast A von T gibt mit

$$\mathbf{M} \models A\sigma.$$

Dieser von dem üblichen[1] abweichende Erfüllbarkeitsbegriff wird zuerst in [Hähnle & Schmitt, 1991] eingeführt, allerdings für die Prädikatenlogik ohne Gleichheit. Mit seiner Hilfe kann die Korrektheit der δ–Regel aus Definition 2.1.3 bewiesen werden.

Schließlich noch eine Definition, die wichtig ist für den Beweis der Vollständigkeit des Tableaukalküls:

Definition 2.1.19 (Faire Tableau–Konstruktionsregel)

Eine Tableau–Konstruktionsregel legt fest, welche von mehreren möglichen Erweiterungsregeln als nächste auf ein Tableau anzuwenden ist. Sie stützt sich dabei auf das Tableau selbst, sowie auf eine Menge von zusätzlichen Informationen. Die Regel gibt an, wie diese zusätzlichen Informationen aufzubauen sind.

Eine Tableau–Konstruktionsregel $\mathcal{R}$ heißt fair, wenn jedes Tableau T_n in der Folge der Tableaus $T_1, T_2, \ldots$ für eine Formel X, die gemäß $\mathcal{R}$ konstruiert wurde, folgende Eigenschaften hat

1. *Auf jede Formel in T_n, die kein Atom ist, wird in jedem Ast, in dem sie auftritt, die auf sie passende Tableau–Erweiterungsregel schließlich angewendet.*
2. *Auf jede γ–Formel in T_n wird in jedem Ast, in dem sie auftritt, die γ–Regel beliebig oft angewendet.*

Ein Beispiel für eine faire Tableau–Konstruktionsregel ist in [Fitting, 1990], Seite 178, angegeben. Die Existenz fairer Konstruktionsregeln ist also sichergestellt.

2.2 Anforderungen an einen Tableaukalkül

Ein Tableaukalkül für die Prädikatenlogik mit Gleichheit muß — wie jeder Tableaukalkül — aus semantischer Sicht gewissen Anforderungen genügen. Dies wird deutlich, wenn man den Aufbau eines Tableaus für eine Formel X als Konstruktion der möglichen Modelle dieser Formel auffaßt.

Denn dafür muß der Kalkül folgendes leisten:

1. Alle Formeln, die in einem Modell $\mathbf{M}$ von X gelten, müssen auf den mit $\mathbf{M}$ korrespondierenden Ast eines Tableaus gebracht werden, und
2. Alle Formeln oder Formelmengen, die in einem Modell nicht gelten können, müssen erkannt werden.

In den kanonischen Modellen, die man für die Prädikatenlogik mit Gleichheit betrachtet, müssen nun zum einen mehr Formeln Gültigkeit haben als in einem beliebigen Modell der Prädikatenlogik (gilt beispielsweise in einem Modell $P(a)$ und $a \approx b$, so muß in einem kanonischen Modell auch $P(b)$ gelten).

Zum anderen gibt es zu einem kanonischen Modell zusätzliche Formeln, die nicht gelten können, etwa $\neg(a \approx a)$.

Um beide Fälle abzudecken, müssen in einem Tableaukalkül für die Prädikatenlogik mit Gleichheit zusätzliche Regeln für die Erweiterung eines Tableaus eingeführt werden und auch zusätzliche Möglichkeiten, einen Ast abzuschließen.

[1] Gewöhnlich, etwa in [Fitting, 1990], wird die Erfüllbarkeit eines Tableaus in der Weise definiert, daß zu jeder Grundsubstitution σ die Existenz einer Interpretation $\mathbf{M}_\sigma$ gefordert wird, die Modell eines Astes des Tableaus ist. In Definition 2.1.18 dagegen wird die Existenz *einer* Interpretation $\mathbf{M}$ gefordert, die die gewünschte Eigenschaft unabhängig von der gewählten Grundsubstitution σ hat.

In der Literatur sind hierfür verschiedene Verfahren angegeben, die in den folgenden Abschnitten beschrieben werden. Wie später deutlich wird, besteht auch die Möglichkeit, die Konstruktion der zusätzlich gültigen Formeln und damit einen Teil der Problematik, einen Tableaubeweis zu finden, in den Abschluß eines Astes zu verlagern. Dies hat, wie sich zeigen wird, gewisse Vorteile.

2.3 Ansätze in der Literatur

2.3.1 Der Ansatz von Jeffrey

Der folgende Ansatz ist der einfachste und intuitiv naheliegendste für die Behandlung der Gleichheit im Tableaukalkül. Er ist beschrieben in [Jeffrey, 1967], kürzer und zusammenfassend auch in [Reeves, 1987]. Alle weiteren Verfahren bauen mehr oder weniger auf ihm auf.

Jeffreys Ansatz gründet sich auf einen Tableaukalkül ohne freie Variablen. Es werden also die γ– und die δ–Regel aus Definition 2.1.2 verwendet.

Wie in Abschnitt 2.2 beschrieben, muß eine Möglichkeit gegeben sein, Formeln, die in einem *kanonischen* Modell zusätzlich gelten müssen, zu einem Ast hinzuzufügen.

Hierzu führt Jeffrey als zusätzliche Tableauerweiterungsregel ein, daß die Gleichungen auf einem Ast des Tableaus in beide Richtungen auf Terme des Astes angewendet werden dürfen, um neue Formeln zu generieren:

Definition 2.3.1 (Tableauerweiterungsregel von Jeffrey)

Sei T ein Tableau für X. Ein Ast A von T enthalte eine Gleichung $\mathsf{T}\,(t \approx s)$ und eine Formel Z, die den Term t (bzw. s) enthält.

Die Formel Z' erhalte man dadurch, daß genau ein Vorkommen von t (bzw. s) in Z durch s (bzw. t) ersetzt wird.

Dann ist das Tableau T', das dadurch entsteht, daß Z' an A angefügt wird, wieder ein Tableau für X.

Schematisch:

$$\begin{array}{c} \mathsf{T}\,(t \approx s) \\ Z(t) \\ \hline Z(s) \end{array} \qquad\qquad \begin{array}{c} \mathsf{T}\,(t \approx s) \\ Z(s) \\ \hline Z(t) \end{array}$$

Beispiel 2.3.2

Abbildung 2.1 zeigt ein Beispiel für eine Tableauerweiterung nach Jeffrey. Man beachte, daß $\mathsf{T}\,P(b,b)$ nicht direkt, sondern erst über den Zwischenschritt $\mathsf{T}\,P(a,b)$ abgeleitet werden kann.

Neben der Erweiterungsregel gibt es eine zusätzliche Abschlußregel. Ein Ast ist auch weiterhin geschlossen, wenn er Formeln $\mathsf{F}\,G$ und $\mathsf{T}\,G$ enthält. Zusätzlich gilt aber:

Definition 2.3.3 (Abschlußregel von Jeffrey)

Ein Ast eines Tableaus, der eine Formel

$$\mathsf{F}\,(t \approx t)$$

enthält, ist geschlossen.

Nach Reeves ([Reeves, 1987]) ist Jeffreys Ansatz „useless in practice“, da die zusätzliche Erweiterungsregel symmetrisch ist, Gleichungen also in beide Richtungen angewendet werden können. Diese Symmetrie führt tatsächlich zu einem sehr großen Suchraum. Enthält nämlich beispielsweise ein Ast die Formeln $f(a) \approx a$ und $P(a)$, so können die Formeln $P(f(a))$, $P(f(f(a)))$, $P(f(f(f(a))))$, usw. zu dem Ast hinzugefügt werden. Diese Möglichkeit würde entfallen, wenn die Erweiterungsregel nicht symmetrisch wäre.

$$\begin{array}{c} \vdots \\ | \\ \mathsf{T}\,(a \approx b) \\ | \\ \mathsf{T}\,P(a,a) \\ | \\ \vdots \end{array} \quad \rightsquigarrow \quad \begin{array}{c} \vdots \\ | \\ \mathsf{T}\,(a \approx b) \\ | \\ \mathsf{T}\,P(a,a) \\ | \\ \mathsf{T}\,P(a,b) \\ | \\ \vdots \end{array} \quad \rightsquigarrow \quad \begin{array}{c} \vdots \\ | \\ \mathsf{T}\,(a \approx b) \\ | \\ \mathsf{T}\,P(a,a) \\ | \\ \mathsf{T}\,P(a,b) \\ | \\ \mathsf{T}\,P(b,b) \\ | \\ \vdots \end{array}$$

Abbildung 2.1: Beispiel für die Erweiterung eines Tableaus nach Jeffrey

2.3.2 Der Ansatz von Reeves

Reeves ([Reeves, 1987]) führt eine zusätzliche Erweiterungsregel ein, die auf der Tatsache beruht, daß in einem kanonischen Modell, in dem Formeln $\mathsf{T}\,P(a_1,\ldots,a_n)$ und $\mathsf{F}\,P(b_1,\ldots,b_n)$ gelten, wenigstens eine der Formeln $\mathsf{F}\,(a_1 \approx b_1)$, …, $\mathsf{F}\,(a_n \approx b_n)$ gelten muß.

Definition 2.3.4 (Tableauerweiterungsregel von Reeves)

Sei T ein Tableau für X. Ein Ast A von T enthalte die potentiell A abschließenden Formeln

1. $\mathsf{T}\,P(a_1,\ldots,a_n)$ *und* $\mathsf{F}\,P(b_1,\ldots,b_n)$, *oder*
2. $\mathsf{F}\,(f(a_1,\ldots,a_n) \approx f(b_1,\ldots,b_n))$.

Dann ist das Tableau T', das man dadurch erhält, daß man die Formel

$$\mathsf{F}\,(a_1 \approx b_1 \wedge \ldots \wedge a_n \approx b_n)$$

an A anfügt, wieder ein Tableau für X. Dabei kann P auch das Gleichheitssymbol $\approx$ sein.

Schematisch:

$$\frac{\mathsf{T}\,P(a_1,\ldots,a_n) \quad \mathsf{F}\,P(b_1,\ldots,b_n)}{\mathsf{F}\,(a_1 \approx b_1 \wedge \ldots \wedge a_n \approx b_n)} \qquad \frac{\mathsf{F}\,(f(a_1,\ldots,a_n) \approx f(b_1,\ldots,b_n))}{\mathsf{F}\,(a_1 \approx b_1 \wedge \ldots \wedge a_n \approx b_n)}$$

Beispiel 2.3.5

Die Korrektheit von Reeves Regel soll an einem kleinen Beispiel verdeutlicht werden. Enthält ein Ast A Formeln $\mathsf{T}\,P(a,b)$ und $\mathsf{F}\,P(a',b')$, so kann die Formel $\mathsf{F}\,(a \approx a' \wedge b \approx b')$ zu dem Ast hinzugefügt werden.

Gelingt es nun, A mit Hilfe dieser neuen Formel abzuschließen, so hat man gezeigt, daß in jedem kanonischen Modell, in dem auch die übrigen Formeln auf A gelten, die Gleichungen $\mathsf{T}\,(a \approx a')$ und $\mathsf{T}\,(b \approx b')$ gültig sind. Dann müssen aber die beiden Formeln $\mathsf{T}\,P(a,b)$ und $\mathsf{F}\,P(a',b')$ mit dem gleichen Wahrheitswert belegt sein. Dies ist nicht möglich, und ein Widerspruch ist gezeigt.

In gewissem Sinne wird durch Reeves Regel das mit einem Ast assoziierte Modell genauer konstruiert als beim Ansatz von Jeffrey. Enthält nämlich ein Ast zwei Formeln $\mathsf{T}\,P(f(a))$ und $\mathsf{F}\,P(f(b))$, die in dem Modell gültig sein müssen, so werden hier die Formeln $\mathsf{F}\,(f(a) \approx f(b))$ und $\mathsf{F}\,(a \approx b)$ zu dem Ast hinzugefügt.

Dies führt dazu, daß der Abschluß eines Astes mit zwei komplementären Formeln nicht mehr benötigt wird. Der bei Jeffrey nur als zusätzliche Regel eingeführte Abschluß durch eine Formel $\mathsf{F}\,(t \approx t)$ genügt, um die Vollständigkeit des Kalküls sicherzustellen. Allerdings nur, wenn neben der Regel aus Definition 2.3.4 auch die Symmetrie von Gleichungen zur Tableauerweiterung zugelassen

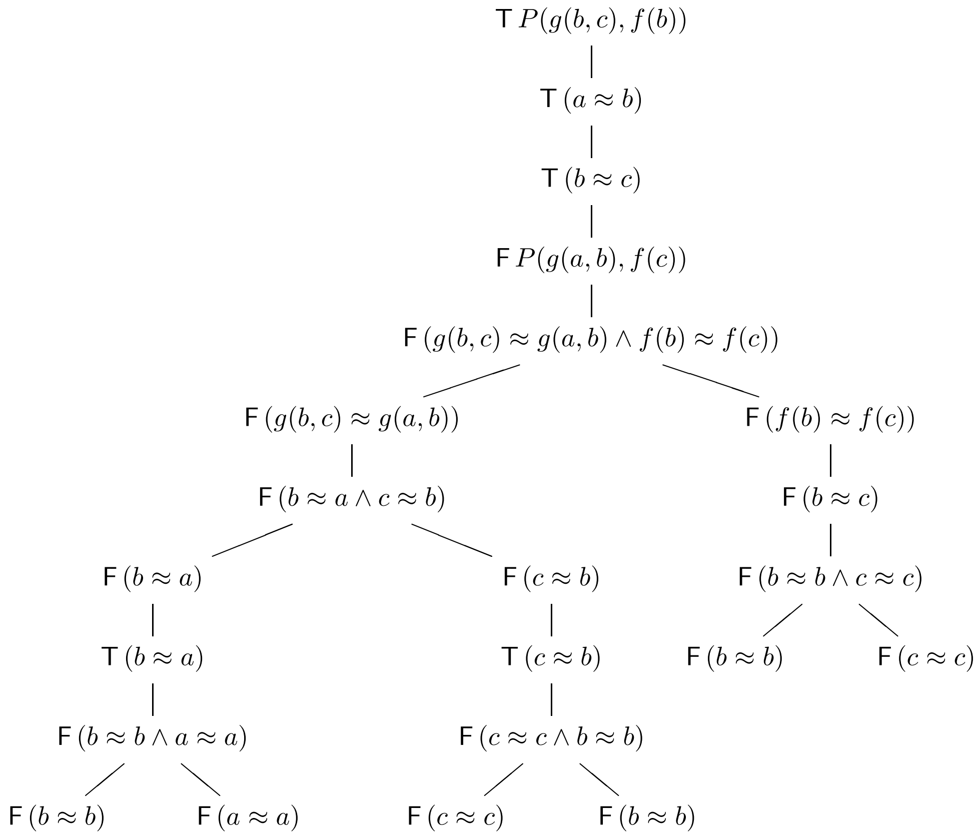


Abbildung 2.2: Ein Tableaubeweis nach Reeves (Beispiel 2.3.6).

ist. Ein Ast, der eine Gleichung T $(t \approx s)$ enthält, muß also um die Gleichung T $(s \approx t)$ erweitert werden können.

Beispiel 2.3.6

Abbildung 2.2 zeigt ein Beispiel für das Verfahren von Reeves. Das nach Definition 2.3.4 aufgebaute Tableau ist ein Beweis für

$$\frac{\begin{array}{c} P(g(b,c), f(b)) \\ a \approx b \\ b \approx c \end{array}}{P(g(a,b), f(c))}$$

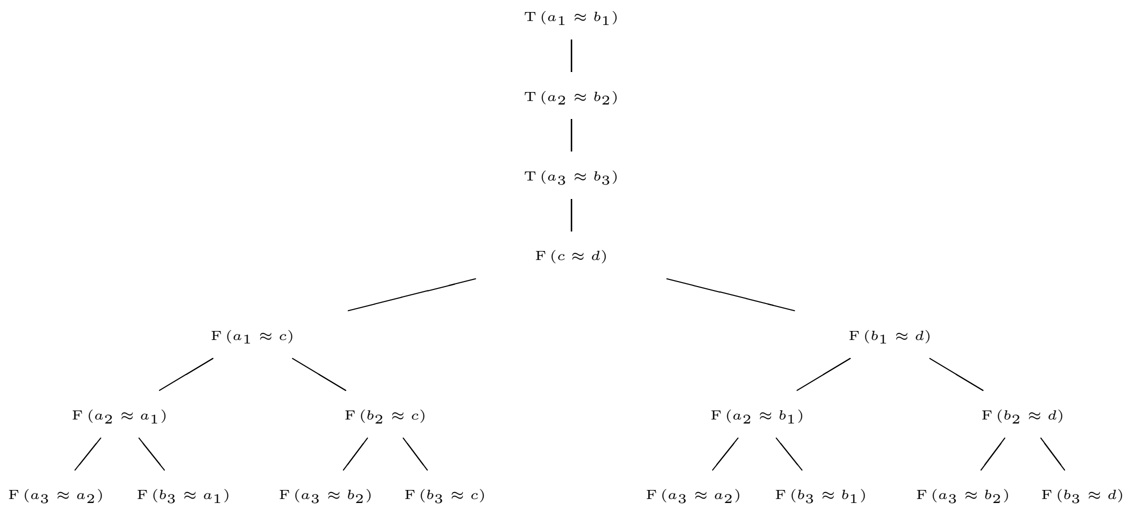
Das Verfahren von Reeves hat sicher Vorteile gegenüber Jeffreys Ansatz, da die Suche nach einem Tableaubeweis eher gerichtet und der Suchraum verkleinert ist.

In einer Beziehung hat das Verfahren jedoch einen entscheidenden Nachteil, der dazu führt, daß es sich nicht für eine Implementierung eignet. Enthält nämlich ein Ast mehrere Gleichungen und eine Ungleichung, tritt das Phänomen auf, das Abbildung 2.3 zeigt. Es kommt zu einer nicht mehr beherrschbaren Menge von Verzweigungen, deren Zahl exponentiell mit der Zahl der Gleichungen auf dem Ast wächst, — es sei denn man verfügt über eine Heuristik, die zum Beweis nicht notwendige Tableauerweiterungen vermeidet.

2.3.3 Der Ansatz von Popplestone

Popplestones Verfahren ([Popplestone, 1967]) stellt im wesentlichen eine Verbesserung des Ansatzes von Jeffrey im Hinblick auf eine mögliche Implementierung dar.

Es werden die gleichen Tableauerweiterungsregeln und Regeln für den Abschluß eines Astes wie bei Jeffreys Kalkül verwendet, insbesondere werden also auch hier keine freien Variablen zugelassen.



Dieses Tableau baut auf drei Gleichungen und einer Ungleichung auf. Ohne weiteres entstehen daraus acht Äste. Dies sind aber nicht alle, das Tableau kann noch weiter verzweigen.

Abbildung 2.3: Nachteil des Verfahrens von Reeves

Die wichtige Neuerung ist, daß an jede Formel im Tableau ein Graph geheftet wird, der die gesamte bekannte Ableitbarkeitsinformation über die Terme enthält, die oberhalb des Knotens im Tableau auftreten oder aus solchen Termen ableitbar sind. Die Terme bilden die Knoten der Graphen. Zwei Terme sind durch eine Kante verbunden, wenn bekannt ist, daß ihre Gleichheit mit den oberhalb von ihnen im Tableau auftretenden Gleichungen gezeigt werden kann.

Tritt nun eine Formel $\mathsf{F}\,(t \approx s)$ in einem Ast auf, und ist diese Formel mit einem Graphen markiert, in dem ein Pfad von t nach s besteht, so ist ein Widerspruch gezeigt und der Ast geschlossen.

Dieses Verfahren hat mehrere Vorteile. Zum einen können beliebige Heuristiken eingesetzt werden, um die Suche in den Graphen und deren Erweiterung zu steuern. Zum anderen wird vermieden, daß die Ableitbarkeitsinformation zu einem Term, also die Information darüber, welche anderen Terme aus ihm ableitbar sind, mehrfach erzeugt wird.

Aber auch Popplestones Ansatz hat noch Nachteile. So wird die Ableitbarkeitsinformation zu Formeln und den Termen die sie enthalten erzeugt, die erkennbar nicht an einem Abschluß beteiligt sein können.

Außerdem hätte eine Implementierung dieses Verfahrens einen massiven Eingriff in die gesamte Struktur der schon bestehenden Module von $_3T^AP$ bedeutet, da die Anwendung der Gleichungen, die ja in der Erweiterung der Graphen besteht, nicht von der Anwendung der übrigen Tableauerweiterungsregeln getrennt ist.

2.3.4 Der Ansatz von Fitting

Fitting beschreibt in seinem Buch ([Fitting, 1990]) zwei leicht unterschiedliche Ansätze, beide jeweils in einer Version mit und einer ohne die Verwendung freier Variablen.

Einer von Fittings Ansätzen stimmt mit dem Jeffreys überein. Fitting favorisiert jedoch sein anderes Verfahren, das sich auf eine nicht symmetrische Erweiterungsregel stützt: eine Gleichung darf nur in eine Richtung angewendet werden. Dadurch geht zunächst die Vollständigkeit des Kalküls verloren. Sie wird dadurch wiederhergestellt, daß an jeden Ast für jeden Term t die Gleichung $\mathsf{T}\,(t \approx t)$ angefügt werden darf.

Definition 2.3.7 (Tableauerweiterungsregel von Fitting)

Sei T ein Tableau für X. Ein Ast A von T enthalte eine Gleichung $\mathsf{T}\,(t \approx s)$ und eine Formel Z, die den Term t enthält.

Die Formel Z' entstehe dadurch, daß man genau ein Vorkommen von t in Z durch s ersetzt.

Dann ist das Tableau T', das man dadurch erhält, daß man Z' an A anfügt, wieder ein Tableau für X.

*Ebenso ist das Tableau T'', das dadurch entsteht, daß für ein $t \in$ **Term** die Formel $\mathsf{T}\,(t \approx t)$ an A angefügt, wird ein Tableau für X.*

Schematisch:

$$\frac{\begin{array}{c}\mathsf{T}\,(t \approx s)\\ Z(t)\end{array}}{Z(s)} \qquad\qquad \frac{}{\mathsf{T}\,(t \approx t)}$$

Auf die Einführung einer zusätzlichen Abschlußregel verzichtet Fitting hier ganz. Tritt nämlich eine Ungleichung $\mathsf{F}\,(t \approx t)$ auf einem Ast auf, kann die komplementäre Gleichung $\mathsf{T}\,(t \approx t)$ ohne weiteres hinzugefügt werden, um den Ast nach der üblichen Regel abzuschließen.

Fitting begründet die Einführung einer nicht symmetrischen Erweiterungsregel damit, daß hierdurch der Suchraum eingeschränkt werde. Auf den ersten Blick ist dies sicherlich richtig. Es ist jedoch möglich, die Umkehrung einer Gleichung $\mathsf{T}\,(a \approx b)$, die auf einem Ast liegt, zu diesem Ast hinzuzufügen: Nach 2.3.7 kann man den Ast um $\mathsf{T}\,(a \approx a)$ erweitern. Auf die linke Seite dieser neuen Gleichung wendet man die ursprüngliche Gleichung $\mathsf{T}\,(a \approx b)$ an und erhält $\mathsf{T}\,(b \approx a)$.

Mit der Möglichkeit, eine Gleichung umzukehren, entartet die nicht symmetrische Erweiterungsregel zu einer Heuristik. Eine Gleichung wird *zunächst* nur in eine Richtung angewendet. Sie kann aber jederzeit umgekehrt und danach auch in die andere Richtung angewendet werden. Dann kann man aber auch bei der symmetrischen Regel bleiben, und es tatsächlich einer Heuristik überlassen, die Richtung zu bestimmen, in der eine Gleichung angewendet wird.

Aus diesem Grunde ist hier die Erweiterung um die Verwendung freier Variablen nur für die symmetrische, schon von Jeffrey eingeführte, Regel wiedergegeben.

Durch die Anwendung der Tableauerweiterungsregel mit freien Variablen wird ein Tableau nicht nur erweitert, sondern insgesamt verändert. Soll nämlich eine Gleichung angewendet werden, und muß dafür eine freie Variable durch einen Term substituiert werden, so muß diese Substitution auf das ganze Tableau erstreckt werden.

Definition 2.3.8 (MGU–Tableauerweiterungsregel von Fitting)

Sei T ein Tableau für X. Ein Ast A von T enthalte eine Gleichung $\mathsf{T}\,(t \approx s)$ und eine Formel Z, in der ein Term t' (bzw. ein Term s') auftritt, der mit t (bzw. mit s) unifizierbar ist. σ sei ein allgemeinster Unifikator (MGU) von t und t' (bzw. von s und s').

Dann ist das Tableau T' wieder ein Tableau für X, das man dadurch erhält, daß man ein Vorkommen von $t'\sigma$ in $Z\sigma$ durch $s\sigma$ ersetzt (bzw. ein Vorkommen von $s'\sigma$ in $Z\sigma$ durch $t\sigma$), und die auf diese Weise neu entstandene Formel an den Ast $A\sigma$ von $T\sigma$ anfügt.

Schematisch:

$$\frac{\begin{array}{c}\mathsf{T}\,(t \approx s)\\ Z(t')\\ \sigma \textit{ MGU von } t \textit{ und } t'\end{array}}{\begin{array}{c}Z(s)\sigma\\ \sigma \textit{ ist auf das ganze Tableau anzuwenden}\end{array}}$$

$$\frac{\begin{array}{c}\mathsf{T}\,(t \approx s)\\ Z(s')\\ \sigma \textit{ MGU von } s \textit{ und } s'\end{array}}{\begin{array}{c}Z(t)\sigma\\ \sigma \textit{ ist auf das ganze Tableau anzuwenden}\end{array}}$$

Unabhängig von der Erweiterung um die Behandlung von Gleichheit führt Fitting eine modifizierte Abschlußregel ein, die auf der Verwendung freier Variablen basiert. Auch hier muß eine den Abschluß zulassende Substitution auf das ganze Tableau angewendet werden.

Definition 2.3.9 (Abschluß nach Fitting mit freien Variablen)

Sei T ein Tableau für eine Formel X. Ein Ast A von T enthalte Formeln $\mathsf{T}\,Z$ und $\mathsf{F}\,Z'$. σ sei ein MGU von Z und Z'.

Dann ist $T\sigma$ ebenfalls ein Tableau für X, und der Ast $A\sigma$ von $T\sigma$ ist geschlossen.

Beispiel 2.3.10

Abbildung 2.4 zeigt ein Beispiel für einen Tableaubeweis mit freien Variablen und Gleichheit. Dieses Beispiel beinhaltet die wesentlichen Probleme bei der Suche eines Beweises in einem Tableaukalkül mit Gleichheit: die Auswahl der richtigen Gleichungsanwendung und der damit verbundenen Variablensubstitution. In Abbildung 2.4 ist die günstigste Reihenfolge von Gleichungsanwendungen und damit ein minimaler Tableaubeweis gezeigt.

Das Beispieltableau stellt einen Beweis dar für

$$\begin{array}{c} (\forall x)(\neg(x \approx a) \vee (g(x) \approx f(x))) \\ (\forall x)(g(f(x)) \approx x) \\ b \approx c \\ P(g(g(a)), b) \\ \hline P(a, c) \end{array}$$

Es ist mit der Substitution $\{x_1 \leftarrow a,\, x_2 \leftarrow a\}$ geschlossen.

Nach Definition 2.3.8 sind die für die beiden Gleichungsanwendungen notwendigen Substitutionen $\{x_1 \leftarrow a\}$ und $\{x_2 \leftarrow a\}$ — wie in Abbildung 2.4 angedeutet[2] — auf das ganze Tableau anzuwenden.

Verwendet man die Erweiterungs– und die Abschlußregeln für freie Variablen, kommt der Reihenfolge, in der Gleichungen angewendet und Äste geschlossen werden, eine große Bedeutung zu. Eine Regelanwendung kann die Anwendung einer anderen Regel oder einen Abschluß verhindern, da das Tableau insgesamt durch die auszuführende Substitution modifiziert wird.

Dennoch bedeutet die Verwendung freier Variablen einen großen Fortschritt. Zwar stellt die Wahl günstiger Regelanwendungen und der mit ihnen verbundenen auszuführenden Substitutionen ein großes Problem dar, verwendet man aber keine freien Variablen, so stellt sich dieses Problem schon bei der Anwendung der γ–Regel, da man schon dann entscheiden muß, welchen Term man für die allquantifizierte Variable einsetzt. Im Vergleich dazu ist der Suchraum durch die Verwendung freier Variablen erheblich eingeschränkt — weswegen sie bei $_3T^AP$ nicht nur bei der Behandlung der Gleichheit eingesetzt werden.

Aus dem bisher in diesem Abschnitt beschriebenen Verfahren entwickelte Fitting eine Implementierung in PROLOG, die in [Fitting, 1990] vollständig wiedergegeben ist.

Um den Kalkül zu implementieren muß eine weitere Problematik gelöst werden: der bisher noch in den Regeln für den Aufbau und Abschluß eines Tableaus enthaltene Indeterminismus muß beseitigt werden, und zwar in einer Weise, daß die Vollständigkeit erhalten bleibt.

Hierzu führt Fitting eine wesentliche Neuerung ein: Die Anwendung der Gleichungen ist nicht in die Anwendung der übrigen Tableauerweiterungsregeln integriert. Zuerst wird die Anwendung der anderen Regeln erschöpft. Auch die γ-Regel wird in dem Sinne erschöpft, daß sie mit einer gewissen Häufigkeit angewendet wird und dann nicht mehr.

Fittings Beweiser läuft sozusagen in zwei Modi, dem normalen Tableauerweiterungs–Modus und dem Gleichheits–Modus, in dem versucht wird, mit Hilfe der Gleichungen einen Abschluß zu finden. Der Vorteil ist, daß wegen der Trennung der Modi Gleichungen nur noch auf atomare Formeln angewendet werden müssen. Dadurch wird

1. der Suchraum verkleinert, und die Zahl der möglichen Gleichungsanwendungen stark eingeschränkt,

[2] x_1/a bedeutet etwa, daß an dieser Stelle im Tableau zunächst die freie Variable x_1 auftritt, diese aber beim weiteren Aufbau des Tableaus durch die Konstante a ersetzt wird, nämlich bei der Anwendung einer Substitution auf das ganze Tableau.

(1) $\mathsf{T}\,(\forall x)(\neg(x \approx a) \vee (g(x) \approx f(x)))$

(2) $\mathsf{T}\,(\forall x)(g(f(x)) \approx x)$

(3) $\mathsf{T}\,(b \approx c)$

(4) $\mathsf{T}\,P(g(g(a)), b)$

(5) $\mathsf{F}\,P(a, c)$

(6) $\mathsf{T}\,(g(f(x_1/a)) \approx x_1/a)$

(7) $\mathsf{T}\,(\neg(x_2/a \approx a) \vee (g(x_2/a) \approx f(x_2/a)))$

(8) $\mathsf{T}\,\neg(x_2/a \approx a)$ (9) $\mathsf{T}\,((g(x_2/a) \approx f(x_2/a))$

(10) $\mathsf{F}\,(x_2/a \approx a)$ (11) $\mathsf{T}\,P(g(f(a)), b)$

(12) $\mathsf{T}\,P(a, b)$

(13) $\mathsf{T}\,P(a, c)$

Das Tableau ist folgendermaßen aufgebaut: Zunächst erhält man die Formeln (6) bis (10) in gewohnter Weise durch Anwendung der Tableau–Erweiterungsregeln aus den Formeln (1) und (2). Danach kann der linke Ast des Tableaus durch Anwendung der Substitution $\{x_2 \leftarrow a\}$ geschlossen werden. Der rechte Ast wird dann mit Hilfe der Gleichungen (9), (6) und (3) um die Formeln (11), (12) und (13) erweitert. Dabei muß, um die Gleichung (6) auf (11) anwenden zu können, die Substitution $\{x_1 \leftarrow a\}$ ausgeführt werden. Die Formel (13) bildet schließlich zusammen mit der Formel (5) den Abschluß des rechten Astes.

Abbildung 2.4: Beispiel für einen Tableaubeweis nach Fitting mit freien Variablen und Gleichheit (Beispiel 2.3.10).

2. es möglich, sich auf solche atomare Formeln zu beschränken, die potentiell den Ast abschließen, zu denen also ein anderes Atom existiert mit gleichem Prädikatensymbol und komplementärer Signatur.

Zwar kann die Trennung der beiden Modi auch von Nachteil sein, nämlich dann, wenn ein Ast durch die Anwendung von Gleichungen auf komplexe Formeln geschlossen werden könnte, bevor durch die Anwendung von Tableauregeln auf diese komplexen Formeln der Ast verzweigt. In der Praxis sind diese Fälle aber selten. Auch ist die Anwendung von Gleichungen sehr viel aufwendiger als die der normalen Tableauerweiterungsregeln.

Durch die Einführung einer Grenze q für die Anwendung der γ–Regel geht die Vollständigkeit des Kalküls verloren. Der Satz über die Vollständigkeit muß daher in der Weise umformuliert werden, daß ein geschlossenes Tableau für eine unerfüllbare Formel gefunden wird, vorausgesetzt, die Grenze q ist hinreichend hoch gewählt.

Die Behebung der nach der Trennung noch verbleibenden Indeterminismen, insbesondere die Auswahl eines von mehreren möglichen Abschlüssen eines Astes oder einer von mehreren möglichen Gleichungsanwendungen, die zu verschiedenen Substitutionen für die freien Variablen führen, wird bei Fittings Implementierung PROLOGs Backtracking überlassen. Alle Möglichkeiten werden also der Reihe nach ausprobiert. Dies führt dazu, daß Fittings Beweiser sehr ineffizient ist — alle anderen denkbaren Verfahren hätten aber sicherlich den Rahmen seines Buches gesprengt. Dies ist wohl der Grund für seinen Schlußsatz ([Fitting, 1990], Seite 231):

> "But remember the resulting system, while complete, may take years to prove anything interesting. Good heuristics are vital now."

2.4 Knuth–Bendix–Verfahren

Das Knuth–Bendix–Verfahren oder ein ähnlicher Ansatz ist für den Einsatz im Tableaukalkül prinzipiell geeignet, es müssen jedoch einige Voraussetzungen erfüllt sein, da im Allgemeinen nicht alle auftretenden Gleichungen bezüglich der Variablen, die sie enthalten, universell sind.

Ein dem Knuth–Bendix–Algorithmus ähnliches Verfahren, das im Tableaukalkül eingesetzt werden könnte, müßte neben den Gleichungen, die eine Gleichungs–Theorie vervollständigen, zu jeder dieser neuen Gleichungen die Information liefern, wie die neuen Gleichungen durch alte ersetzt werden können.

Mit dieser Information könnte man dann nachdem man festgestellt hat, daß zwei Terme unter der Annahme des All–Abschlusses gleich sind, überprüfen, ob diese Gleichheit auch mit den nicht all–abgeschlossenen Gleichungen gezeigt werden kann, und welche Möglichkeiten es dafür gibt, wie also die nicht all–abgeschlossenen Variablen instantiiert werden müssen. So könnten dann schließlich alle möglichen Abschlüsse eines Astes generiert werden.

Es scheint allerdings nicht einfach zu sein, das Knuth–Bendix–Verfahren entsprechend zu erweitern. Auch würden die bekannten Probleme der Termgewichtung zur Orientierung der Gleichungen auftreten.

Dennoch kann das Knuth–Bendix–Verfahren, insbesondere wenn man die im folgenden vorgestellten Ergebnisse berücksichtigt, gewinnbringend eingesetzt werden. Da sich dies jedoch erst aufgrund der vorliegenden Untersuchung zeigte, wurde auf die Verwendung von Knuth–Bendix–Algorithmen zunächst verzichtet. In Abschnitt 6.3 wird kurz skizziert, auf welche Weise das Knuth–Bendix–Verfahren in den vorgestellten Rahmen eingepaßt werden könnte.

Kapitel 3

Universelle Formeln

3.1 Theorie universeller Formeln

Im folgenden ist das Konzept der in einem Tableau bezüglich einer Variablen universellen Formeln beschrieben. Es wurde zunächst nur für Gleichungen im Tableau entwickelt, dann aber auf beliebige Formeln erweitert, so daß es nicht mehr ganz in den durch das Thema dieser Studienarbeit, die Gleichheit im Tableaukalkül, gesteckten Rahmen paßt. Darum ist der Beschreibung dieses Konzeptes ein eigenes Kapitel gewidmet.

Alle Definitionen und Sätze diese Kapitels beziehen sich auf nicht signierte Formeln. Um sie auf ein Tableau mit Vorzeichen zu erweitern, kann die Korrespondenz zwischen signierten und nicht signierten Formeln aus Definition 2.1.4 verwendet werden.

Eines der größten Probleme bei der Suche nach einem Beweis im Tableaukalkül, also eines geschlossenen Tableaus, ist die Frage, wann und wie oft die γ–Regel angewendet werden soll.

Dieses Problem ist besonders schwerwiegend in der praktischen Anwendung, wenn Theoreme einer durch Axiome beschriebenen mathematischen Theorie bewiesen werden sollen.

Als Beispiel soll hier die Gruppentheorie dienen. Das die Assoziativität beschreibende Axiom

$$(\forall x)(\forall y)(\forall z)[(x \cdot y) \cdot z \approx x \cdot (y \cdot z)]$$

ist für das auftretende Problem typisch. Selbst zum Beweis einfachster Sätze, wie beispielsweise der Eindeutigkeit des neutralen Elements in einer Gruppe, muß die Assoziativität schon etwa fünf– bis zehnmal mit verschiedenen Instantiierungen der Variablen eingesetzt werden.

In einem Tableaubeweis muß dafür die γ–Regel entsprechend oft angewendet werden. Da man aber nicht weiß, welche γ–Formel oft verwendet werden muß, und durch die häufige Anwendung der γ–Regel auch Verzweigungen entstehen können, ist es wünschenswert, Formeln von der Art des Assoziativitäts–Axioms zu erkennen und dann ohne mehrfache Anwendung der γ–Regel mit verschiedenen Instantiierungen ihrer freien Variablen zum Abschluß von Ästen verwenden zu können.

Die Formeln der oben beschriebenen Form sind in gewissem Sinne universell bezüglich der (oder einiger der) Variablen, die sie enthalten. Diese Tatsache spiegelt sich in der folgenden Definition der universellen Formeln wieder, die sie aus modelltheoretischer Sicht charakterisiert:

Definition 3.1.1 (Universelle Formel)

Eine Formel F, die auf dem Ast A eines Tableaus liegt, heißt universell bezüglich einer Variablen x, wenn für jedes Modell $\mathbf{M}$ und jede Grundsubstitution σ gilt:

Wenn $\mathbf{M} \models A\sigma$, dann $\mathbf{M} \models ((\forall x)F)\sigma$.

Dies gilt insbesondere immer dann, wenn x in F nicht frei vorkommt.

Um die Eigenschaft der Universalität für einen Tableaubeweis ausnutzen zu können, braucht man ein einfaches Verfahren, um universelle Formeln zu erkennen. Ein solches Verfahren existiert für einen Teil der universelle Formeln. Zunächst legt die folgende Definition aber fest, welche Anforderungen ein solches Verfahren erfüllen muß.

Definition 3.1.2 (Methode zum Auffinden universeller Formeln)

Eine Abbildung

$$\mathcal{M} : \{A \;:\; A \textit{ der Ast eines Tableaus}\} \times \mathbf{Form} \longrightarrow 2^{\mathbf{Var}}$$

heißt Methode zum Auffinden universeller Formeln, wenn aus

$$x \in \mathcal{M}(A, F)$$

folgt, daß die Formel F auf A liegt, und universell ist bezüglich x.

Dieser Definition genügt beispielsweise die triviale Methode

$$\mathcal{M}_0 = \emptyset,$$

die überhaupt keine universellee Formel auffindet.

Leistungsfähiger aber immer noch hinreichend einfach, so daß sie in der Praxis einsetzbar ist, ist die im folgenden Satz eingeführte Methode.

Satz 3.1.3

$\mathcal{M}(A, F)$ enthalte genau die Variablen x, für die gilt: F liegt auf A und ist

1. *mit Hilfe der γ–Regel aus einer γ–Formel abgeleitet, und x ist die freie Variable, die dabei für die durch den Quantor gebundene Variable eingesetzt wurde, oder*
2. *mit Hilfe der α–Regel oder der γ–Regel aus einer Formel abgeleitet, die universell ist bezüglich x.*

Dann ist $\mathcal{M}$ eine Methode zum Auffinden universeller Formeln.

Beweis:

F liege auf A, und für ein beliebiges Modell $\mathbf{M}$ und eine Substitution σ gelte:

$$\mathbf{M} \models A\sigma.$$

Es gelte $x \in \mathcal{M}(A, F)$, also trifft einer der beiden folgenden Fälle zu:

1. (a) $F = F(x)$ ist aus $(\forall y)F(y)$ abgeleitet.
Dann gilt $(\forall y)F(y) \in A$, damit

$$\mathbf{M} \models ((\forall y)F(y))\sigma \in A\sigma$$

und also

$$\mathbf{M} \models ((\forall x)F(x))\sigma.$$

1. (b) $F = \neg G(x)$ ist aus $\neg(\exists y)G(y)$ abgeleitet.
Dann kann man analog zu Fall 1 (a) folgern: $\neg(\exists y)G(y) \in A$ und daher

$$\mathbf{M} \models (\neg(\exists y)G(y))\sigma \in A\sigma.$$

Dann gilt aber wegen $\{\neg(\exists y)G(y)\} \models (\forall x)\neg G(x)$ auch

$$\mathbf{M} \models ((\forall x)\neg G)\sigma \in A\sigma$$

und $F = \neg G(x)$ ist universell bezüglich x.

2. (a) O.B.d.A. $F = \alpha_1 \in A$ und $\alpha \in A$ ist universell bezüglich x.
Dann gilt

$$\mathbf{M} \models ((\forall x)\alpha)\sigma$$

und da $\{\alpha\} \models \alpha_1 \wedge \alpha_2$ auch

$$\mathbf{M} \models ((\forall x)\alpha_1)\sigma,$$

und $F = \alpha_1$ ist universell bezüglich x.

2. (b) $F = \gamma(y) \in A$ und $\gamma \in A$ ist universell bezüglich x.
Dann gilt

$$\mathbf{M} \models ((\forall x)\gamma)\sigma$$

und also auch

$$\mathbf{M} \models ((\forall x)\gamma(y))\sigma$$

und $F = \gamma(y)$ ist universell bezüglich x. ■

Durch die Methode aus Satz 3.1.3 werden aber bei weitem nicht alle universellen Formeln erfaßt. So ist beispielsweise zu beachten, daß in einem Ast, der für alle Grundsubstitutionen unerfüllbar ist, alle Formeln universell bezüglich aller freien Variablen sind.

Die in Satz 3.1.3 genannten sind aber die einzigen universellen Formeln, die „einfach" zu erkennen sind — wenn auch noch Verfahren denkbar sind, die einige andere Spezialfälle abdecken.

Nachdem festgelegt ist, was unter universellen Formeln zu verstehen ist, und es auch möglich ist, einige von ihnen im Tableau aufzufinden, kann man daran gehen, die Regel für den Abschluß eines Tableaus abzuschwächen. Dadurch werden zum einen die Tableaubeweise verkürzt, zum anderen wird aber auch, was noch wichtiger ist, die Suche nach ihnen vereinfacht.

Zuvor aber sei noch einmal die übliche Regel für den Abschluß eines Tableaus genannt. Sie lautet:

Definition 3.1.4 (Abschluß eines Tableaus)

Ein Tableau ist geschlossen, wenn eine Grundsubstitution σ existiert, so daß jeder Ast des Tableaus Formeln F und $\neg G$ enthält mit

$$F\sigma = G\sigma.$$

Diese Regel kann in der folgenden Weise abgeschwächt werden:

Definition 3.1.5 (Abschluß mit universellen Formeln)

Ein Tableau mit den Ästen $A_1, \ldots, A_k$ ist geschlossen, wenn Grundsubstitutionen $\sigma_1, \ldots, \sigma_k$ existieren, so daß

1. *jeder Ast A_i Formeln F_i und $\neg G_i$ enthält mit*

   $$F_i\sigma_i = G_i\sigma_i,$$

   und

2. *für je zwei Substitutionen σ_i und σ_j $(1 \leq i < j \leq k)$ und alle freien Variablen x gilt: Wenn*

   $$x\sigma_i \neq x\sigma_j,$$

   dann sind für den Abschluß von A_i oder für den Abschluß von A_j nur bezüglich x universelle Formeln verwendet worden, d. h.

   $$F_i \text{ und } G_i \quad \text{oder} \quad F_j \text{ und } G_j$$

   sind universell bezüglich x.

Die Vollständigkeit des Tableaukalküls wird durch die neue Abschlußregel nicht berührt, da sie eine Abschwächung der Standard–Abschlußregel ist. Der folgende Satz aber muß bewiesen werden, der sicher stellt, daß die obige Abschlußregel die Korrektheit erhält.

Satz 3.1.6

Ein nach der Abschlußregel aus Definition 3.1.5 geschlossenes Tableau ist nicht erfüllbar.

Beweis:

Sei das Tableau T mit den Ästen $A_1, \ldots, A_k$ nach Definition 3.1.5 mit den Grundsubstitutionen $\sigma_1, \ldots, \sigma_k$ geschlossen.

Um zu zeigen, daß T nicht erfüllbar ist, ist eine Substitution σ für die freien Variablen anzugeben, so daß keiner der Äste von $T\sigma$ ein Modell hat. Die Substitution σ kann als Kombination von $\sigma_1, \ldots, \sigma_k$ in der folgenden Weise konstruiert werden:

Für jede freie Variable x sei

$$x\sigma = \begin{cases} x\sigma_j & \text{falls für } j \in \{1, \ldots, k\} \text{ der Abschluß von } A_j \text{ nicht} \\ & \text{universell ist bezüglich } x \\ x\sigma_1 & \text{sonst} \end{cases}$$

Die Wohldefiniertheit von σ folgt daraus, daß, wenn $x\sigma_{j_1} \neq x\sigma_{j_2}$ ist, aufgrund der Definition des Abschlusses entweder der Abschluß von A_{j_1} oder der Abschluß von A_{j_2} universell sein muß bezüglich x. Es gibt also immer nur ein j, für das der erste Teil der Definition von σ greift. Der zweite Teil ist nur der Vollständigkeit halber hinzugefügt für den Fall, daß die Abschlüsse aller Äste universell sind bezüglich x.

Nun ist noch zu zeigen, daß keiner der Äste von $T\sigma$ ein Modell hat.

Annahme: $\mathbf{M} \models A_i\sigma$ $(1 \leq i \leq k)$
Der Ast A_i enthält Formeln F_i und $\neg G_i$ mit $F_i\sigma_i = G_i\sigma_i$. Aus der Annahme folgt, daß $\mathbf{M} \models F_i\sigma, \neg G_i\sigma$.

Um zu einem Widerspruch zu kommen, genügt es nun zu zeigen, daß daraus $\mathbf{M} \models F_i\sigma_i, \neg G_i\sigma_i$ folgt, da $F_i\sigma_i$ und $\neg G_i\sigma_i$ komplementäre Formeln sind.

Seien $\{x_1, \ldots, x_p, y_1, \ldots, y_q\}$ die in F_i und G_i frei auftretenden Variablen und $x_1, \ldots, x_p$ die Variablen bezüglich derer F_i und G_i universell sind.

Aus der Definition von σ folgt, daß

$$y_m\sigma = y_m\sigma_i \quad (1 \leq m \leq q).$$

Aus $\mathbf{M} \models F_i\sigma, \neg G_i\sigma$ folgt mit der Definition der universellen Formeln

$$\mathbf{M} \models ((\forall x_1) \ldots (\forall x_p)F_i)\sigma,\ ((\forall x_1) \ldots (\forall x_p)\neg G_i)\sigma.$$

In den Formeln $(\forall x_1) \ldots (\forall x_p)F_i$ und $(\forall x_1) \ldots (\forall x_p)\neg G_i$ kommen nur noch die Variablen $y_1, \ldots, y_q$ frei vor. Auf diesen stimmen σ und σ_i überein. Daher gilt

$$\mathbf{M} \models ((\forall x_1) \ldots (\forall x_p)F_i)\sigma_i,\ ((\forall x_1) \ldots (\forall x_p)\neg G_i)\sigma_i,$$

woraus sofort

$$\mathbf{M} \models F_i\sigma_i, \neg G_i\sigma_i$$

folgt, was zum Widerspruch führt. ■

Beispiel 3.1.7

Abbildung 3.1 zeigt ein einfaches Beispiel für den Abschluß eines Tableaus mit Hilfe der neuen Regel. In diesem Beispiel ist der linke Ast mit der Substitution $\sigma_1 = \{x \leftarrow a\}$ und der rechte Ast mit der Substitution $\sigma_2 = \{x \leftarrow b\}$ geschlossen. σ_1 und σ_2 unterscheiden sich zwar in der Belegung der Variablen x, dies ist aber zulässig, da alle am Abschluß der Äste beteiligten Formeln, nämlich $p(x)$, $\neg p(a)$ und $\neg p(b)$, nach Satz 3.1.3 universell sind bezüglich x.

3.2 Praktische Anwendung des Konzeptes der universellen Formeln

Um ein prädikatenlogisches Tableau mit freien Variablen nach der normalen Abschlußregel aus Definition 3.1.4 zu schließen, müssen nicht nur komplementäre Formeln auf den Ästen gefunden werden, sondern auch eine Substitution, die es erlaubt, alle Äste simultan zu schließen.

Die Suche nach einer solchen Substitution kann dadurch geschehen, daß die Äste der Reihe nach geschlossen werden, und jeweils die für den Abschluß notwendigen Variablenbelegungen ausgeführt

$$\neg((\forall x)p(x) \supset (p(a) \land p(b)))$$
$$|$$
$$(\forall x)p(x)$$
$$|$$
$$\neg(p(a) \land p(b))$$
$$|$$
$$p(x)$$
$$\diagup \qquad \diagdown$$
$$\neg p(a) \qquad\qquad \neg p(b)$$

Abbildung 3.1: Ein nach Definition 3.1.5 geschlossenes Tableau (Beispiel 3.1.7).

werden. Diese sammeln sich dann zu der gewünschten Substitution an, wenn es gelingt alle Äste zu schließen.

Ist dies aber nicht möglich, so muß man durch Backtracking und erneute Behandlung schon geschlossener Äste andere Abschlüsse mit anderen Variablenbelegungen finden. Dies ist das in $_{3}T^{A}P$ angewandte Verfahren.

Die Suche nach einem Abschluß mit der neuen Regel gestaltet sich etwas einfacher. Zwar müssen wiederum komplementäre Formeln auf allen Ästen gefunden werden. Es muß nun aber nicht eine Substitution angegeben werden für das ganze Tableau, sondern eine für jeden Ast. Nach der Definition des Abschlusses (Def. 3.1.5) darf sich die einem Ast A zugeordnete Substitution in der Belegung einer Variablen x von allen anderen Substitutionen unterscheiden, wenn beide am Abschluß des Astes A beteiligten Formeln universell sind bezüglich der Variablen x.

Um diese Substitutionen zu finden, kann man genauso vorgehen, wie bei der Suche nach einem normalen Abschluß — mit dem Unterschied, daß nicht, wie es dort der Fall ist, eine Substitution durch Instantiieren aller an einem Abschluß beteiligten freien Variablen konstruiert wird. Statt dessen wird eine Variable x nur dann belegt, wenn der Abschluß des Astes auf Formeln beruht, die bezüglich x nicht universell sind. Dann nämlich ist die Belegung von x für die weiteren zu findenden Substitutionen bindend, es sei denn, daß zur Bildung dieser weiteren Substitutionen auf die Instantiierung von x verzichtet werden kann.

Daß das neue Verfahren tatsächlich günstiger ist als das Standard–Verfahren, liegt daran, daß weniger Variablen instantiiert werden, es also seltener vorkommt, daß durch eine schon ausgeführte Belegung der Abschluß eines weiteren Astes verhindert wird.

3.3 Universelle Formeln aus Sicht des Tableau–Aufbaus

Neben dem formalen Beweis für die Korrektheit der neuen Abschlußregel, der hauptsächlich modelltheoretisch argumentiert, kann man auch eine einfache Begründung aus syntaktischer, auf die Suche nach einem Abschluß gerichteter Sicht geben:

Ist eine Formel F auf einem Ast als universell bezüglich einer Variablen x mit Hilfe von Satz 3.1.3 erkannt, also mit Hilfe der dort genannten Regeln abgeleitet worden, so ist es jederzeit möglich, eine neue Formel $F' = F\{x \leftarrow u\}$ auf dem Ast zu erzeugen, wobei u im Tableau bisher nicht auftritt.

Darum ist es, wenn man einen Abschluß mit der Formel F macht, nicht notwendig, x zu belegen, da man statt dessen auch den Abschluß mit F' machen könnte, wodurch an Stelle von x die sonst nirgends auftretende Variable u belegt würde.

Dies gilt allerdings nicht für beliebige universelle Formeln, sondern tatsächlich nur für die nach Satz 3.1.3 als universell erkannten.

Kapitel 4

Konzeption eines Verfahrens für die Behandlung der Gleichheit

4.1 Aus der Literatur übernommene Anregungen

Das im folgenden beschriebene Konzept für die Behandlung der Gleichheit stützt sich nicht völlig auf einen der in der Literatur beschriebenen und in Kapitel 2.3 wiedergegebenen Ansätze. Vielmehr sind aus jedem dieser Ansätze einige Ideen übernommen und mit einer Reihe eigener Überlegungen zu einem neuen Konzept verschmolzen worden:

- Aus [Fitting, 1990]: Trennung der Anwendung von Gleichungen von der Anwendung der übrigen Tableau–Erweiterungsregeln,
- Aus [Popplestone, 1967]: für jeden Term wird die Ableitbarkeitsinformation soweit möglich nur einmal erzeugt,
- Aus [Reeves, 1987]: Paare potentiell abschließender Atome werden in Disjunktionen von Ungleichungen überführt, die durch Anwendung der Gleichungen zum Widerspruch geführt werden.

Die zusätzlichen eigenen Überlegungen — sie werden in den folgenden Abschnitten beschrieben — sind vor allem darauf gerichtet, das Verfahren so zu gestalten, daß es nicht nur vollständig und korrekt ist, sondern in der Anwendung tatsächlich zu Ergebnissen führt.

4.2 Das zugrunde gelegte Verfahren

Den Ausgangspunkt für alle weiteren Überlegungen bildet das in diesem Abschnitt beschriebene Verfahren, ein Tableaukalkül mit freien Variablen und Gleichheit. Es ist eine Abwandlung des von Fitting zur Implementierung vorgeschlagenen Verfahrens (Abschn. 2.3.4). In den weiteren Abschnitten wird es verbessert und zu dem schließlich im Rahmen dieser Arbeit implementierten Kalkül erweitert.

Das wesentliche Merkmal des die Grundlage bildenden Kalküls ist die Trennung der Erweiterung des Tableaus durch eine faire Konstruktionsregel von der Anwendung der Gleichheit und des Abschlusses des Tableaus.

Der Aufbau eines Tableaus ist in den folgenden Definitionen beschrieben:

Definition 4.2.1 (Einfache Formel)

Eine Formel heißt einfach, wenn sie

1. *ein positiv signiertes Atom (auch eine Gleichung) ist oder*

2. *eine Ungleichung, d. h. von der Form* $\mathsf{F}\,(t \approx s)$.

Definition 4.2.2 (Tableau mit freien Variablen und Gleichheit)

Sei $q \geq 1$ *eine Grenze für die Anwendung der* γ*–Regel, und*

$\mathcal{R}$ *sei eine unter Beachtung der Grenze* q *für die Anwendung der* γ*–Regel faire Tableaukonstruktionsregel, die die Tableauerweiterungsregeln mit freien Variablen (Def. 2.1.1 und 2.1.3) verwendet.*

Ein Tableau ist in folgender Weise aufgebaut:

1. *Zunächst werden gemäß* $\mathcal{R}$ *Erweiterungsregeln angewendet, bis das Tableau unter Beachtung der Grenze* q *voll expandiert ist.*
2. *Dann wird eine Grundsubstitution* σ *ausgewählt und auf das Tableau angewandt.*
3. *Schließlich wird das nun keine freien Variablen mehr enthaltende Tableau durch beliebig häufige Anwendung der symmetrischen Gleichheitsregel (Def. 2.3.1) auf einfache Formeln erweitert.*

Man muß beachten, daß nur der erste Teil des Tableauaufbaus nach Definition 4.2.2 deterministisch verläuft. Die anzuwendende Grundsubstitution σ und die Reihenfolge der Gleichungensanwendungen können frei gewählt werden.

Definition 4.2.3 (Abschluß eines Tableaus mit freien Variablen und Gleichheit)

Ein nach Definition 4.2.2 aufgebautes Tableau ist geschlossen, wenn jeder seiner Äste

1. *atomare Formeln* $\mathsf{T}\,P(t_1,\ldots,t_n)$ *und* $\mathsf{F}\,P(t_1,\ldots,t_n)$ *oder*
2. *eine Ungleichung* $\mathsf{F}\,(t \approx t)$

enthält.

Beispiel 4.2.4

Die Abbildungen 4.1 und 4.2 zeigen ein Beispiel für ein nach Definition 4.2.2 aufgebautes Tableau.

Wie das Tableau aus Beispiel 2.3.10 stellt es einen Beweis dar für

$$\begin{array}{c} (\forall x)(\neg(x \approx a) \vee (g(x) \approx f(x))) \\ (\forall x)(g(f(x)) \approx x) \\ b \approx c \\ P(g(g(a)), b) \\ \hline P(a,c) \end{array}$$

Im Gegensatz zu Fittings Verfahren wird hier die Substitution

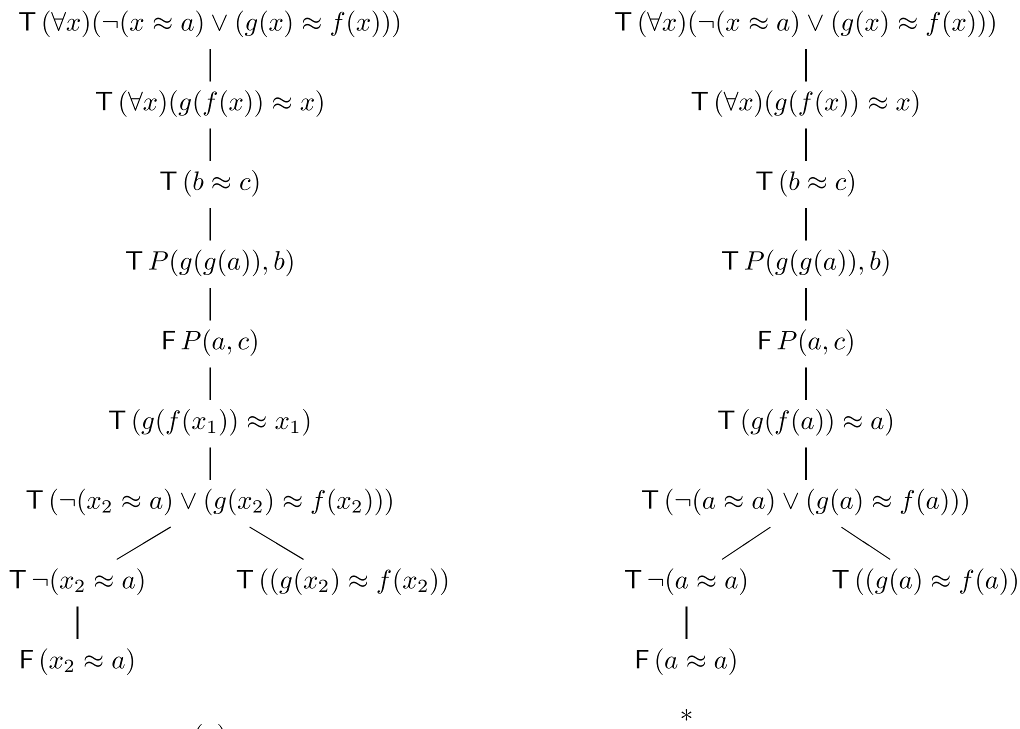
$$\sigma = \{x_1 \leftarrow a,\ x_2 \leftarrow a\}$$

nicht zur Anwendung einer Gleichung oder zum Abschluß eines Astes ausgeführt — also bei Bedarf —, sondern nachdem das Tableau unter Beachtung der Grenze $q = 1$ für die Anwendung der γ–Regel voll expandiert ist (Abb. 4.1 (a)). In Abbildung 4.1 (b) ist das Tableau nach Anwendung der Substitution gezeigt.

Anschließend wird das Tableau (sein rechter Ast) durch Gleichungsanwendungen erweitert, bis das in Abbildung 4.2 gezeigte Tableau entsteht. Es ist nach Definition 4.2.3 geschlossen. Der linke Ast enthält die Ungleichung $\mathsf{F}\,(a \approx a)$, der rechte Ast die komplementären Atome $\mathsf{T}\,P(a,c)$ und $\mathsf{F}\,P(a,c)$.

Man beachte, daß die durchaus naheliegende Substitution $\{x_2 \leftarrow g(a)\}$ den Abschluß beider Äste verhindern würde.

Zum Beweis der Korrektheit des Kalküls werden die folgenden beiden Lemmata benötigt:



(a) (b)

Abbildung 4.1: Die ersten beiden Schritte des Aufbaus eines Tableaus nach Definition 4.2.2 (Beispiel 4.2.4).

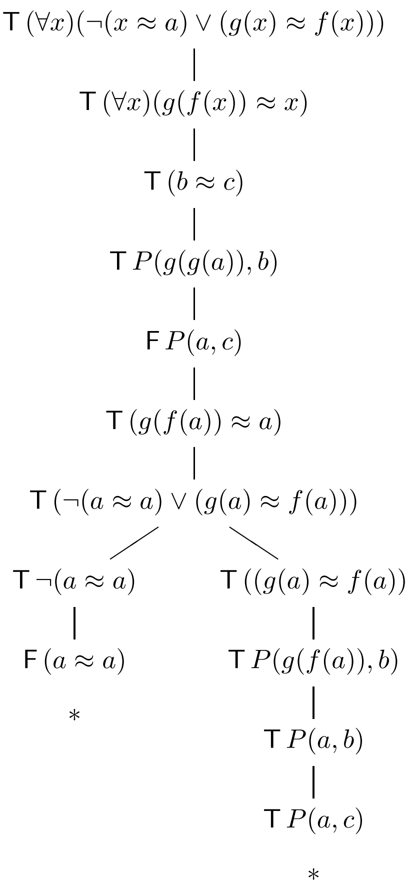


Abbildung 4.2: Der dritte Schritt des Aufbaus eines Tableaus nach Definition 4.2.2 (Beispiel 4.2.4).

Lemma 4.2.5

Ein nach Definition 4.2.2 aufgebautes Tableau für eine erfüllbare Formel X ist erfüllbar im Sinne von Definition 2.1.18.

Beweis:

Das initiale Tableau, mit dem der Aufbau beginnt, ist erfüllbar, da es nur die erfüllbare Formel X enthält.

Es genügt also zu zeigen, daß jede Erweiterung nach Definition 4.2.2 die Erfüllbarkeit eines Tableaus erhält.

Für die im ersten Schritt des Aufbaus angewandten normalen Tableauerweiterungsregeln, insbesondere die neue δ–Regel, ist dies in [Hähnle & Schmitt, 1991] bewiesen (für beliebige Modelle; der Beweis läßt sich aber völlig analog auf die hier betrachteten kanonischen Modelle übertragen).

Daß der zweite Schritt, die Anwendung der Grundsubstitution, die Erfüllbarkeit erhält, folgt sofort aus deren Definition (Def. 2.1.18).

Nachdem der zweite Schritt des Tableauaufbaus ausgeführt ist, enthält das Tableau keine freien Variablen mehr. Daß die Erfüllbarkeit eines solchen Tableaus durch Anwendung der symmetrischen Regel für die Gleichheit nicht verloren geht, ist nun leicht zu zeigen:

Die auf dem Ast A eines Tableaus T liegende Gleichung $\mathsf{T}\,(t \approx s)$ sei auf eine Formel $Z(t)$ angewendet und die Formel $Z(s)$ zu A hinzugefügt worden, das Ergebnis sei das Tableau T'.

Das keine freien Variablen mehr enthaltende Tableau T sei erfüllbar. Dann gibt es eine kanonische Interpretation $\mathbf{M}$, die Modell eines Astes von T ist. Ist dies ein anderer als der Ast A, so ist dieser andere Ast A' auch in T' enthalten und T' daher erfüllbar. Gilt andernfalls $\mathbf{M} \models A$, dann gilt insbesondere

$$\mathbf{M} \models Z(t),\, t \approx s.$$

Dann gilt aber auch

$$\mathbf{M} \models Z(s)$$

und daher

$$\mathbf{M} \models A'$$

Damit ist T' wieder erfüllbar nach Definition 2.1.18. ■

Lemma 4.2.6

Ein nach Definition 4.2.2 aufgebautes und nach Definition 4.2.3 geschlossenes Tableau ist nicht erfüllbar.

Beweis:

Annahme: Es gibt ein geschlossenes, erfüllbares Tableau T
Wegen seiner Erfüllbarkeit enthält T einen Ast A, der ein kanonisches Modell hat.

Dies ist aber nicht möglich, da auch der Ast A, wie alle Äste von T, entweder zwei komplementäre Atome oder eine Ungleichung identischer Terme enthält. In keinem kanonischen Modell können aber zwei komplementäre Atome zugleich oder eine Ungleichung identischer Terme gelten.

Die Annahme ist also falsch, und ein geschlossenes Tableau ist nicht erfüllbar. ■

Zum Beweis der Vollständigkeit des Kalküls werden der Begriff der Hintikka–Menge und Hintikkas Lemma benötigt:

Definition 4.2.7 (Hintikka–Menge)

Eine Menge $\mathcal{H}$ geschlossener Formeln heißt Hintikka–Menge der Prädikatenlogik mit Gleichheit, wenn sie folgende Bedingungen erfüllt:

1. *Wenn* $\mathsf{T}\,\neg X \in \mathcal{H}$*, dann* $\mathsf{F}\,X \in \mathcal{H}$*.*
2. *Wenn* $\mathsf{F}\,\neg X \in \mathcal{H}$*, dann* $\mathsf{T}\,X \in \mathcal{H}$*.*
3. *Wenn* $\alpha \in \mathcal{H}$*, dann* $\alpha_1 \in \mathcal{H}$ *und* $\alpha_2 \in \mathcal{H}$*.*
4. *Wenn* $\beta \in \mathcal{H}$*, dann* $\beta_1 \in \mathcal{H}$ *oder* $\beta_2 \in \mathcal{H}$*.*
5. *Wenn* $\gamma \in \mathcal{H}$*, dann* $\gamma(t) \in \mathcal{H}$ *für jeden geschlossenen Term* t*.*
6. *Wenn* $\delta \in \mathcal{H}$*, dann gibt es einen geschlossenen Term* t*, so daß* $\delta(t) \in \mathcal{H}$*.*
7. *Wenn*
   - $\mathsf{T}\,(t \approx s) \in \mathcal{H}$ *oder* $\mathsf{T}\,(s \approx t) \in \mathcal{H}$*,*
   - $X(t) \in \mathcal{H}$ *und*
   - $X(t)$ *einfach,*

   dann $X(s) \in \mathcal{H}$*.*
8. $\mathsf{F}\,(t \approx t) \notin \mathcal{H}$ *für jeden geschlossenen Term* t*.*

Lemma 4.2.8

Ist die Menge $\mathcal{H}$ *eine Hintikka–Menge der Prädikatenlogik mit Gleichheit, so hat* $\mathcal{H}$ *ein kanonisches Modell.*

Beweis:

Hintikkas Lemma findet sich als Proposition 8.8.3 in [Fitting, 1990] und ist dort bewiesen.
■

Der Beweis der Vollständigkeit des Kalküls wird durch den Indeterminismus, der in der Definition für den Aufbau eines Tableaus (Def. 4.2.2) enthalten ist, erschwert. Dieser Indeterminismus besteht darin, daß die Grenze q für die Anwendung der γ–Regel und die Reihenfolge der Gleichungsanwendungen frei bestimmt werden kann. Vorübergehend soll der Indeterminimus nun beseitigt werden. Dafür wird hilfsweise ein etwas anderer Kalkül definiert und seine Vollständigkeit bewiesen. Darauf wird dann der Beweis der Vollständigkeit des Kalküls nach Definition 4.2.2 und Definition 4.2.3 gestützt.

Da es gerade der Indeterminismus ist, der es erlaubt, kurze Tableaubeweise zu finden, ist das hilfsweise betrachtete Verfahren für die praktische Anwendung völlig ungeeignet. Es wird ausschließlich für den Vollständigkeitsbeweis benötigt, und dann im weiteren nicht mehr verwendet.

Definition 4.2.9 (Hilfsweise definierter Tableaukalkül)

Der hilfsweise definierte Tableaukalkül soll mit dem nach Definition 4.2.2 und Definition 4.2.3 übereinstimmen, mit Ausnahme folgender Punkte:

1. *Es gibt keine Grenze für die Anwendung der* γ*–Regel.*
2. *Es wird eine* γ*–Regel ohne freie Variablen verwendet. Bei der Anwendung der* γ*–Regel wird also sofort ein geschlossener Term für die durch den Quantor gebundene Variable eingesetzt.*

   Die einzusetzenden Terme werden in folgender Weise „fair“ ausgewählt: Ist $t_1, t_2, \ldots$ *eine Aufzählung aller geschlossenen Terme, dann wird bei der ersten Anwendung der* γ*–Regel auf eine bestimmte* γ*–Formel auf einem bestimmten Ast der Term* t_1 *eingesetzt, bei der zweiten Anwendung* t_2 *usw. Dies führt dazu, daß schließlich jeder Term auf jedem Ast für jede* γ*–Formel verwendet wird.*
3. *Es wird eine* δ*–Regel verwendet, die der Regel für freie Variablen entspricht.*

   Bei der Anwendung der δ*–Regel wird also nicht eine neue Skolemkonstante für die gebundene Variable eingesetzt, sondern der mit einer Skolemfunktion aufgebaute Term, der eingesetzt würde, wenn das Tableau mit freien Variablen aufgebaut wäre, und die freien Variablen später mit dem Term belegt würden, der nun aber tatsächlich schon sofort an Stelle der freien Variablen eingesetzt wird.*

   Ein Beispiel für die Anwendung dieser δ*–Regel zeigt Abbildung 4.3.*

$$\begin{array}{c} \mathsf{T}\,(\forall x)(\exists y)P(x,y) \\ | \\ \mathsf{T}\,(\exists y)P(c,y) \\ | \\ \mathsf{T}\,P(c,f(c)) \end{array}$$

Abbildung 4.3: Beispiel für die Anwendung der in Definition 4.2.9, Punkt 3, hilfsweise definierten δ–Regel.

4. *Die Anwendung der Gleichungen wird mit der Anwendung der übrigen Tableauerweiterungsregeln vermischt. Dies ist möglich, da hier zu keiner Zeit freie Variablen auftreten.*
5. *Die Tableaukontruktionsregel $\mathcal{R}'$ nach der das Tableau aufgebaut wird, ist ein Erweiterung der fairen Regel $\mathcal{R}$. Sie legt nun auch fest, wann welche Gleichungen auf welche einfachen Formeln angewendet werden sollen.*

   Dabei ist $\mathcal{R}'$ auch im Bezug auf die Anwendung von Gleichungen fair, d. h. auf jedem Ast wird jede Gleichung auf jede einfache Formel, auf die sie angewendet werden kann, schließlich auch angewendet.

   Eine in diesem Sinne faire Konstruktionsregel läßt sich beispielsweise dadurch angeben, daß eine geeignete Ordnung auf den einfachen Formeln und den Gleichungen und damit auf der Anwendung von Gleichungen definiert wird.
6. *Der Abschluß eines Tableaus ist wie in Definition 4.2.3 definiert, es muß dazu allerdings keine schließende Grundsubstitution angegeben werden, da das Tableau keine freien Variablen enthält.*

Lemma 4.2.10

Der Tableaukalkül nach Definition 4.2.9 ist vollständig, d. h.

ist X eine unerfüllbare Formel, so gibt es ein nach Definition 4.2.9 aufgebautes und geschlossenes Tableau für X.

Beweis:

Über die Vollständigkeit eines ähnlichen Tableaukalküls wird in [Fitting, 1990] ein Satz aufgestellt (Theorem 8.9.5), für den allerdings nur eine Beweisidee angegeben wird. Auf diese stützt sich der nachfolgende Beweis.

Das Tableauverfahren nach Definition 4.2.9 verläuft völlig deterministisch. Alle Schritte beim Aufbau sind durch die Konstruktionsregel $\mathcal{R}'$ genau festgelegt. Die Tableaus für eine Formel X können also folgendermaßen in eindeutiger Weise bezeichnet werden: sei T_1 das initiale Tableau für X und $T_2, T_3, \ldots$ die Folge der Tableaus, die daraus durch Erweiterung gemäß $\mathcal{R}'$ der Reihe nach entstehen.

Zu zeigen ist also, daß es für eine unerfüllbares X in dieser Folge eine Tableau T_n gibt, so daß alle Äste von T_n geschlossen sind.

Äquivalent zu dieser Behauptung ist: Enthalten alle Tableaus $T_1, T_2, \ldots$ für X einen offenen Ast, so ist X erfüllbar.

Dies wird nun bewiesen. Sei also eine Formel X vorgelegt, und jedes der Tableaus $T_1, T_2, \ldots$ für X möge einen offenen Ast enthalten.

Im allgemeinen ist die Folge der T_n unendlich. Man kann sie daher als eine Annäherung an ein unendliches Tableau T betrachten. Dieses Tableau T enthält dann ebenfalls einen offenen Ast.[1]

Die Menge der auf diesem offenen Ast gelegenen Formeln bildet nun eine Hintikka–Menge der Prädikatenlogik mit Gleichheit (Def. 4.2.7), was sofort aus der Tatsache folgt, daß

[1] Das auf Königs Lemma gestützte Standardargument für diese Folgerung findet sich etwa im Beweis zu Theorem 7.8.6 in [Fitting, 1990].

1. der Ast nicht geschlossen ist, und
2. das Tableau mit Hilfe der fairen Kontruktionsregel $\mathcal{R}'$ aufgebaut wurde, was gewährleistet, daß die weiteren Bedingungen für eine Hintikka–Menge erfüllt sind.

Eine Hintikka–Menge besitzt aber nach 4.2.8 ein Modell, das dann auch Modell der auf jedem Ast des Tableaus liegenden Formel X sein muß. X ist also erfüllbar, was zu zeigen war. ■

Damit kann nun der folgende Satz bewiesen werden, der Vollständigkeit und Korrektheit des oben beschriebenen, grundlegenden Kalküls sicherstellt. Auf ihn stützen sich die Beweise für Vollständigkeit und Korrektheit der in den folgenden Abschnitten beschriebenen Erweiterungen.

Satz 4.2.11

Sei T ein nach Definition 4.2.2 unter Beachtung einer bestimmten Grenze q für die Anwendung der γ–Regel aufgebautes Tableau für eine Formel X.

Korrektheit: Ist T nach Definition 4.2.3 geschlossen, dann ist die Formel X unerfüllbar.

Vollständigkeit: Ist die Formel X unerfüllbar und die Grenze q hinreichend hoch gewählt, dann ist T nach Definition 4.2.3 geschlossen.

Beweis:

Korrektheit:
Nachdem die Lemmata 4.2.5 und 4.2.6 schon zur Verfügung stehen, gelingt der Beweis der Korrektheit problemlos.

Es gebe ein geschlossenes Tableau für die Formel X. Dieses ist nach Lemma 4.2.6 nicht erfüllbar. Dann kann aber wegen Lemma 4.2.5 auch X nicht erfüllbar sein.

Vollständigkeit:
Mit Hilfe der Vollständigkeit des hilfsweise definierten Kalküls (Lemma 4.2.10) wird nun gezeigt, daß auch das Verfahren nach den Definitionen 4.2.2 und 4.2.3 vollständig ist.

Sei also eine unerfüllbare Formel X vorgelegt. Dann existiert ein nach dem Hilfsverfahren aufgebautes Tableau T_H für X, das geschlossen ist. Es wird nun gezeigt, daß dann auch ein nach Definition 4.2.2 aufgebautes und nach 4.2.3 geschlossenes Tableau T' für X existiert.

Die Grenze q für die Anwendung der γ–Regel sei mindestens gleich der größten Zahl von Anwendungen der γ–Regel auf eine γ–Formel auf einem Ast von T_H.

T entstehe nun dadurch, daß T_H gemäß der Konstruktionsregel $\mathcal{R}'$ aus Definition 4.2.9 erweitert wird, bis es unter Beachtung der Grenze q voll expandiert ist. T ist wie T_H geschlossen.

Die nach Definition 4.2.3 frei bestimmbare, den Abschluß bewirkende Substitution σ sei in der Weise festgelegt, daß sie die Einsetzung von Termen durch die γ–Regel beim Tableauaufbau nach dem Hilfsverfahren modelliert, also wie folgt:

Die γ–Formeln in T seien mit $\gamma_1, \gamma_2, \ldots, \gamma_k$ bezeichnet, und die freien Variablen mit x_{ij}, wobei die Variable x_{ij} bei der j–ten Anwendung der γ–Regel auf γ_i für die gebundene Variable eingesetzt wird.

Dann sei die Substitution σ bestimmt durch

$$\sigma = \{x_{ij} \leftarrow t_j \, : \, 1 \leq i \leq k, 1 \leq j \leq q\}.$$

Nun ist noch zu bestimmen, in welcher Weise die Gleichungen zum Aufbau von T' angewendet werden sollen, nachdem die Anwendung der normalen Erweiterungsregeln gemäß Definition 4.2.2 abgeschlossen ist.

Dazu muß man sich zunächst verdeutlichen, daß es unerheblich ist, ob die Anwendung von Gleichungen mit der Anwendung der übrigen Regeln vermischt oder bis zum Schluß aufgeschoben wird. Der Grund hierfür ist, daß auf die durch die Anwendung von Gleichungen entstehenden neuen Formeln keine Erweiterungsregeln angewendet werden können, da sie einfach sind. Das Aufschieben der Gleichungsanwendungen führt also nur dazu, daß sich die Reihenfolge der Formeln auf einem Ast ändert, nicht aber die sonstige Form des Tableaus.

Unter Umständen muß allerdings eine Gleichungsanwendung, die vorher oberhalb einer Verzweigung des Tableaus ausgeführt wurde, wenn sie nun erst hinter der Verzweigung liegt, auf beiden Teilästen, also mehrfach ausgeführt werden.

Sei also T' das Tableau, daß dadurch entsteht, daß man alle Gleichungsanwendungen in T bis zum Schluß aufschiebt, das aber auf allen seinen Ästen die gleichen Formeln enthält wie T.

Man muß nun nachvollziehen, daß T' damit tatsächlich gemäß Definition 4.2.2 aufgebaut ist. Teil 1 dieser Definition ist erfüllt, da die zum Aufbau von T verwendete Konstruktionsregel $\mathcal{R}'$ eine Erweiterung von $\mathcal{R}$ ist, und nachdem alle Gleichungsanwendungen an den Schluß verschoben sind, T nach der Regel $\mathcal{R}$ aufgebaut ist. Allerdings enthält T' in diesem Stadium im Gegensatz zu T freie Variablen.

Als Grundsubstitution werde nun σ ausgeführt, womit Teil 2 von Definition 4.2.2 erfüllt ist.

Schließlich ist auch Teil 3 erfüllt, da die an den Schluß verschobene Anwendung der Gleichungen nun, da die Substitution σ angewandt wurde, ausgeführt werden können.

T' ist geschlossen, da alle Äste von T' die gleichen Formeln enthalten, die den Abschluß der entsprechenden Äste von T bewirken.

Damit ist gezeigt, daß es zu einer beliebigen unerfüllbaren Formel X ein nach Definition 4.2.2 unter Beachtung einer hinreichend hohen Grenze q aufgebautes Tableau gibt, daß nach Definition 4.2.3 geschlossen ist. ■

Nachdem die Korrektheit und Vollständigkeit des in diesem Abschnitt beschriebenen Kalküls bewiesen sind, kann nun daran gegangen werden, diesen Kalkül zu verbessern und so zu erweitern, daß er für die Implementierung geeignet ist.

4.3 Berechnung der Gleichungstheorie für potentiell abschließende Atome

Ist der normale Tableaumodus beendet, das Tableau unter Berücksichtigung der Einschränkung bezüglich der γ–Regel also voll expandiert, so kann es — vor oder nach der Anwendung einer Grundsubstitution — in der folgenden Weise in eine andere für die Behandlung der Gleichheit besser geeignete Datenstruktur überführt werden. Diese neue Datenstruktur besteht zum einen aus der Menge der auf einem Ast A gelegenen Gleichungen $\mathrm{Gl}(A)$, und zum anderen aus einer Menge von Disjunktionen von Ungleichungen $\mathrm{Dis}(A)$:

Definition 4.3.1 ($\mathrm{Gl}(A)$ und $\mathrm{Dis}(A)$)

1. *Die Menge* $\mathrm{Gl}(A)$ *ist definiert als*

$$\mathrm{Gl}(A) := \{s \approx t : \mathsf{T}\,(s \approx t) \in A\}.$$

2. *Die Menge* $\mathrm{Dis}(A)$ *enhält*
   - *für jedes Paar potentiell abschließender Atome* $\mathsf{T}\,P(t_1, \ldots, t_n)$ *und* $\mathsf{F}\,P(s_1, \ldots, s_n)$ *auf A die n–stellige Disjunktion*

     $$t_1 \not\approx s_1 \vee \ldots \vee t_n \not\approx s_n$$

     und
   - *für jede Ungleichung* $\mathsf{F}\,(t \approx s)$ *auf A die einstellige Disjunktion*

     $$t \not\approx s.$$

Beispiel 4.3.2

Als Beispiel soll das in Abbildung 4.1 (b) dargestellte, voll expandierte Tableau dienen.

Sein rechter Ast $A_{\mathrm{Bsp}}\sigma$ enthält nur ein Paar komplementärer, potentiell den Ast abschließender Atome: $\mathsf{T}\,P(g(g(a)), b)$ und $\mathsf{F}\,P(a, c)$. Darum ist $\mathrm{Dis}(A_{\mathrm{Bsp}}\sigma)$ einelementig:

$$\mathrm{Dis}(A_{\mathrm{Bsp}}\sigma) = \{g(g(a)) \not\approx a \vee b \not\approx c\}.$$

In $\mathrm{Gl}(A_{\mathrm{Bsp}}\sigma)$ sind die drei Gleichungen des Astes enthalten:

$$\mathrm{Gl}(A_{\mathrm{Bsp}}\sigma) = \{b \approx c,\, g(f(a)) \approx a,\, g(a) \approx f(a)\}.$$

Der linke Ast des Tableaus sei mit A'_{Bsp} bezeichnet. Für ihn gilt:

$$\begin{aligned} \mathrm{Dis}(A'_{\mathrm{Bsp}}\sigma) &= \{a \not\approx a,\, (g(g(a)) \not\approx a \vee b \not\approx c)\} \\ \mathrm{Gl}(A'_{\mathrm{Bsp}}\sigma) &= \{b \approx c,\, g(f(a)) \approx a\} \end{aligned}$$

Um ein Tableau abzuschließen, muß eine Substitution σ angegeben werden, so daß für alle Äste A die n Ungleichungen einer Disjunktion nach Anwendung der Substitution σ mit Hilfe der Gleichungen in $\mathrm{Gl}(A\sigma)$ zum Widerspruch geführt werden können, denn damit ist ein Widerspruch bezüglich der ursprünglichen Atome gezeigt.

Die neue Datenstruktur entspricht dem ersten Verzweigungsschritt beim Verfahren von Reeves. Sie vermeidet aber die in Abschnitt 2.3.2 beschriebenen Nachteile, die durch mehrmalige Anwendung von Reeves Erweiterungsregel auf Paare von Gleichungen und Ungleichungen entstehen.

Außerdem ist die neue Datenstruktur voll abgestimmt auf die Anwendung der Gleichheit. Man kann nun beliebige Algorithmen verwenden, um die Ungleichungen zum Widerspruch zu führen, also die Gleichheit der Terme zu zeigen. Dies kann beispielsweise dadurch geschehen, daß man sukzessive die Äquivalenzklassen der Terme bezüglich der durch die Gleichungen auf dem Ast gegebenen Gleichheitsrelation berechnet:

Definition 4.3.3 (Äquivalenzklasse)

Für einen geschlossenen Term t bezeichne $[t]_A$ die Äquivalenzklasse bezüglich der durch die Gleichungen $\mathrm{Gl}(A)$ *auf dem Ast A gegebenen Gleichungstheorie, also bezüglich der Relation die definiert ist durch*

$$t \sim s \quad \textit{gdw.} \quad \begin{array}{l} t = s \textit{ gilt in der durch die Gleichungen in } \mathrm{Gl}(A). \\ \textit{gegebenen Gleichungstheorie} \end{array}$$

Die Definition für den Abschluß eines Tableaus muß nun an das neue Verfahren angepaßt werden:

Definition 4.3.4 (Abschluß eines Tableaus mit freien Variablen und Gleichheit mit Hilfe von Äquivalenzklassen)

Ein Tableau T ist geschlossen, wenn es eine Grundsubstitution σ gibt, so daß für jeden Ast A von T gilt:

Es gibt eine Disjunktion

$$(t_1\sigma \not\approx s_1\sigma \vee \ldots \vee t_n\sigma \not\approx s_n\sigma) \in \mathrm{Dis}(A\sigma),$$

so daß für $1 \le i \le n$ gilt:

$$[t_i\sigma]_{A\sigma} = [s_i\sigma]_{A\sigma}.$$

Beispiel 4.3.5

Als Beispiel diene wieder das schon in Beispiel 4.3.2 verwendete Tableau, auf das schon die Substitution $\sigma = \{x_1 \leftarrow a,\, x_2 \leftarrow a\}$ angewendet wurde.

Die in $\mathrm{Dis}(A_{\mathrm{Bsp}}\sigma)$ liegende Disjunktion $g(g(a)) \not\approx a \vee b \not\approx c$ bewirkt den Abschluß des rechten Astes, denn es gilt

$$\begin{aligned} [g(g(a))]_{A_{\mathrm{Bsp}}\sigma} &= \\ [a]_{A_{\mathrm{Bsp}}\sigma} &= \{a,\, g(g(a)),\, g(f(a)),\, g(g(g(g(a)))), \\ &\qquad g(g(g(f(a)))),\, g(g(f(g(a)))),\, \ldots\ \} \end{aligned}$$

und

$$[b]_{A_{\mathrm{Bsp}}\sigma} = [c]_{A_{\mathrm{Bsp}}\sigma} = \{b,\, c\}.$$

Die Äquivalenzklasse $[a]_{A_{\mathrm{Bsp}}\sigma}$ ist zwar nicht endlich, sie ist aber vergleichsweise stark eingeschränkt, da die Gleichungen keine freien Variablen mehr enthalten.

Der linke Ast A'_{Bsp} ist geschlossen, denn es gilt $(a \not\approx a) \in \mathrm{Dis}(A'_{\mathrm{Bsp}}\sigma)$ und trivialerweise $[a]_{A'_{\mathrm{Bsp}}\sigma} = [a]_{A'_{\mathrm{Bsp}}\sigma}$.

Der Abschluß nach Definition 4.3.4 schließt auch die Fälle ein, in denen auf einem Ast zwei — auch schon ohne die Anwendung von Gleichungen — komplementäre Atome liegen.

Verwendet man die Abschlußregel aus Definition 4.3.4, muß man beim Aufbau des Tableaus keine Gleichungen mehr anwenden. Auch die Anwendung einer Substitution auf das Tableau ist nun in der Definition für den Abschluß enthalten. Die Schritte zwei und drei in der früheren Definition für den Aufbau eines Tableaus (Def. 4.2.2) können also entfallen:

Definition 4.3.6 (Tableau mit freien Variablen)

Sei $q \geq 1$ eine Grenze für die Anwendung der γ–Regel, und

$\mathcal{R}$ sei eine unter Beachtung der Grenze q faire Tableaukonstruktionsregel, die die Tableauerweiterungsregeln mit freien Variablen (Def. 2.1.1 und 2.1.3) verwendet.

Zum Aufbau eines Tableaus werden gemäß $\mathcal{R}$ Erweiterungsregeln angewendet, bis das Tableau unter Beachtung der Grenze q voll expandiert ist.

Beim Aufbau eines Tableaus nach dieser Definition muß die besondere Bedeutung des $\approx$–Prädikates nicht mehr beachtet werden. Der einzige verbliebene Freiheitsgrad beim Aufbau ist die Wahl der Schranke q. Dieses ist nun genau das Verfahren nach dem Tableaus vom Beweiser $_3T^AP$ auch schon vor dem Einbau der Gleichheitsbehandlung aufgebaut wurden.

Die Methode zum Aufbau eines Tableaus muß damit zur Erweiterung von $_3T^AP$ nicht verändert werden. Die gesamte Gleichheitsbehandlung kann in die Überprüfung des Abschlusses eines Tableaus verlegt werden — die damit entsprechend kompliziert wird.

Daß das Verfahren auch mit der neuen Abschlußregel korrekt und vollständig bleibt, stellt das folgende Lemma sicher. Zu seinem Beweis wird der neue Abschluß auf den schon im letzten Abschnitt definierten und dort als korrekt und vollständig bewiesenen zurückgeführt.

Satz 4.3.7

Sei T das nach Definition 4.3.6 unter Beachtung einer bestimmten Grenze q für die Anwendung der γ–Regel aufgebaute Tableau für eine Formel X.

Korrektheit: Ist T nach Definition 4.3.4 geschlossen, dann ist die Formel X unerfüllbar.

Vollständigkeit: Ist die Formel X unerfüllbar und die Grenze q hinreichend hoch gewählt, dann ist T nach Definition 4.3.4 geschlossen.

Beweis:

Korrektheit:

Sei das Tableau T mit den Ästen $A_1, \ldots, A_k$ nach Definition 4.3.4 geschlossen.

Dann gibt es eine Grundsubstitution σ und für jeden Ast A_i $(1 \leq i \leq k)$ eine Disjunktion

$$D_i = (t^i_1\sigma \not\approx s^i_1\sigma \vee \ldots \vee t^i_{n_i}\sigma \not\approx s^i_{n_i}\sigma) \in \mathrm{Dis}(A_i\sigma),$$

so daß für $1 \leq i \leq k$ und $1 \leq j \leq n_i$ gilt:

$$[t^i_j\sigma]_{A_i\sigma} = [s^i_j\sigma]_{A_i\sigma}.$$

Es gibt also eine Folge von Gleichungen $g^i_{j1}, \ldots, g^i_{jm_{ij}} \in \mathrm{Gl}(A_i\sigma)$ $(m_{ij} \geq 0)$, mit deren Hilfe $s^i_j\sigma$ aus $t^i_j\sigma$ abgeleitet werden kann.

Das Tableau T' werde nun entsprechend Definition 4.2.2 wie folgt aufgebaut:

1. Das Tableau T ist so aufgebaut, wie es dem ersten Schritt eines Aufbaus nach Definition 4.2.2 entspricht.
2. Auf T wird entsprechend dem zweiten Schritt in Definition 4.2.2 die Substitution σ angewendet.
3. Schließlich wird nach Schritt 3 in Definition 4.2.2 jeder der Äste $A_i\sigma$ in folgender Weise durch Anwendung der symmetrischen Gleichheitsregel (Def. 2.3.1) erweitert:
   Ist die Disjunktion D_i aus einer Ungleichung $\mathsf{F}(t^i_1 \approx s^i_1)$ hervorgegangen (dann ist $n_i = 1$), werde auf sie die Folge der Gleichungen $g^i_{1,1}, \ldots, g^i_{1,m_{i,1}}$ in der Weise angewendet, daß der entstehende Ast A'_i des Tableaus T' die Ungleichung $\mathsf{F}(s^i_1 \approx s^i_1)$ enthält.

Ist D_i aus zwei potentiell komplementären Atomen $\mathsf{T}\,P(t_1^i, \ldots, t_{n_i}^i)$ und $\mathsf{F}\,P(s_1^i, \ldots, s_{n_i}^i)$ entstanden, so werde auf das positiv signierte als erstes die Folge der Gleichungen $g_{1,1}^i, \ldots, g_{1,m_{i,1}}^i$ angewendet — es entsteht die Formel $\mathsf{T}\,P(s_1^i, t_2^i, \ldots, t_{n_i}^i)$ —, dann die Gleichungsfolge $g_{2,1}^i, \ldots, g_{2,m_{i,2}}^i$ und so fort bis schließlich nach der Anwendung von $g_{n_i,1}^i, \ldots, g_{n_i,m_{i,n_i}}^i$ der Ast A_i' die Formel $\mathsf{T}\,P(s_1^i, \ldots, s_{n_i}^i)$ enthält.

Das entstandene Tableau T' ist also nach Definition 4.2.2 aufgebaut. Es ist auch nach Definition 4.2.3 geschlossen, denn jeder seiner Äste enthält ein Paar komplementärer Atome oder eine Ungleichung der Form $\mathsf{F}\,(s \approx s)$. Aus Satz 4.2.11 folgt nun die Unerfüllbarkeit der Formel X.

Vollständigkeit:
Sei T' das nach Satz 4.2.11 für ein hinreichend großes q existierende, nach Definition 4.2.3 geschlossene Tableau. Das mit dieser Grenze q aufgebaute Tableau T stimmt mit dem Tableau überein, das nach dem ersten Schritt des Aufbaus von T' entstanden war. σ sei die in Schritt 2 auf T' angewandte Grundsubstitution.

Für jeden der Äste A_i' von T' (bzw. A_i von T) kann nun die folgende Überlegung angestellt werden:

A_i' enthält eine Ungleichung $\mathsf{F}\,(r_1 \approx r_1)$ oder aber komplementäre Atome $\mathsf{T}\,P(r_1, \ldots, r_k)$ und $\mathsf{F}\,P(r_1, \ldots, r_k)$.

Dann muß aber schon A_i eine Ungleichung $\mathsf{F}\,(t_1 \approx s_1)$ oder potentiell komplementäre Atome $\mathsf{T}\,P(t_1, \ldots, t_k)$ und $\mathsf{F}\,P(s_1, \ldots, s_k)$ enthalten, und es muß mit Hilfe der Gleichungen in $\mathrm{Gl}(A_i\sigma)$ möglich sein, für $1 \leq j \leq k$ den Term r_j sowohl aus $t_j\sigma$ als auch aus $s_j\sigma$ abzuleiten.

Dann enthält aber $\mathrm{Dis}(A_i\sigma)$ die Disjunktion $(t_1\sigma \not\approx s_1\sigma \vee \ldots \vee t_k\sigma \not\approx s_k\sigma)$, und für $1 \leq j \leq k$ gilt:

$$[t_j\sigma]_{A_i\sigma} = [s_j\sigma]_{A_i\sigma}.$$

Da dies für alle Äste A_i gilt, ist das Tableau T nach Definition 4.3.4 geschlossen. ■

Der wesentliche Vorteil des auf Äquivalenzklassen beruhenden Verfahrens ist, daß keine unnötige Ableitbarkeitsinformation mehr erzeugt wird: Für all die Atome, die nicht potentiell den Ast abschließen, da zu ihnen kein komplementäres Atom existiert, wird keine Ableitbarkeitsinformation erzeugt. Noch wichtiger aber ist, daß Ableitbarkeitsinformationen nicht mehr mehrfach erzeugt werden, wie es bei den Verfahren der Fall ist, die auf zusätzliche Tableauerweiterungsregeln gegründet sind. Dies soll das folgende Beispiel verdeutlichen:

Beispiel 4.3.8

Abbildung 4.4 zeigt den Ast eines Tableaus, der nach Jeffreys Methode (Abschnitt 2.3.1) erweitert und geschlossen wurde. Zum Abschluß wird die Information, daß c aus a ableitbar ist, zweimal erzeugt. Zusätzlich werden aber auch unnötige Zwischenschritte mehrfach ausgeführt, so daß insgesamt acht Ableitungsschritte nötig sind, um den Ast zu schließen.

In Abbildung 4.5 ist das gleiche Tableau mit Hilfe des auf der Umwandlung in Disjunktionen von Ungleichungen beruhenden Ansatzes geschlossen. Hier werden tatsächlich nur die beiden unbedingt notwendigen Ableitungsschritte ausgeführt.

Man muß beachten, daß dieses Beispiel nicht besonders darauf zugeschnitten ist, die Vorteile des neuen Ansatzes gegenüber dem von Jeffrey hervorzuheben — allenfalls dadurch, daß die beiden Stellen des Prädikates P mit den gleichen Termen belegt wurden.

Tatsächlich ist jedes Verfahren, das darauf beruht, das Tableau durch Gleichungsanwendungen zu erweitern, wie die folgende Überlegung zeigt, im allgemeinen Fall sehr viel schlechter als das neue, das sich auf die Berechnung der Äquivalenzklassen stützt, und in keinem Fall besser:

Gegeben sei der Ast A eines Tableaus, der die beiden potentiell den Ast abschließenden Prädikate $\mathsf{T}\,P(t_1, \ldots, t_k)$ und $\mathsf{F}\,P(s_1, \ldots, s_k)$ enthalte. Um die weiteren Betrachtungen zu vereinfachen enthalte der Ast keine weiteren Paare potentiell abschließender Atome. Auch sei die Problematik,

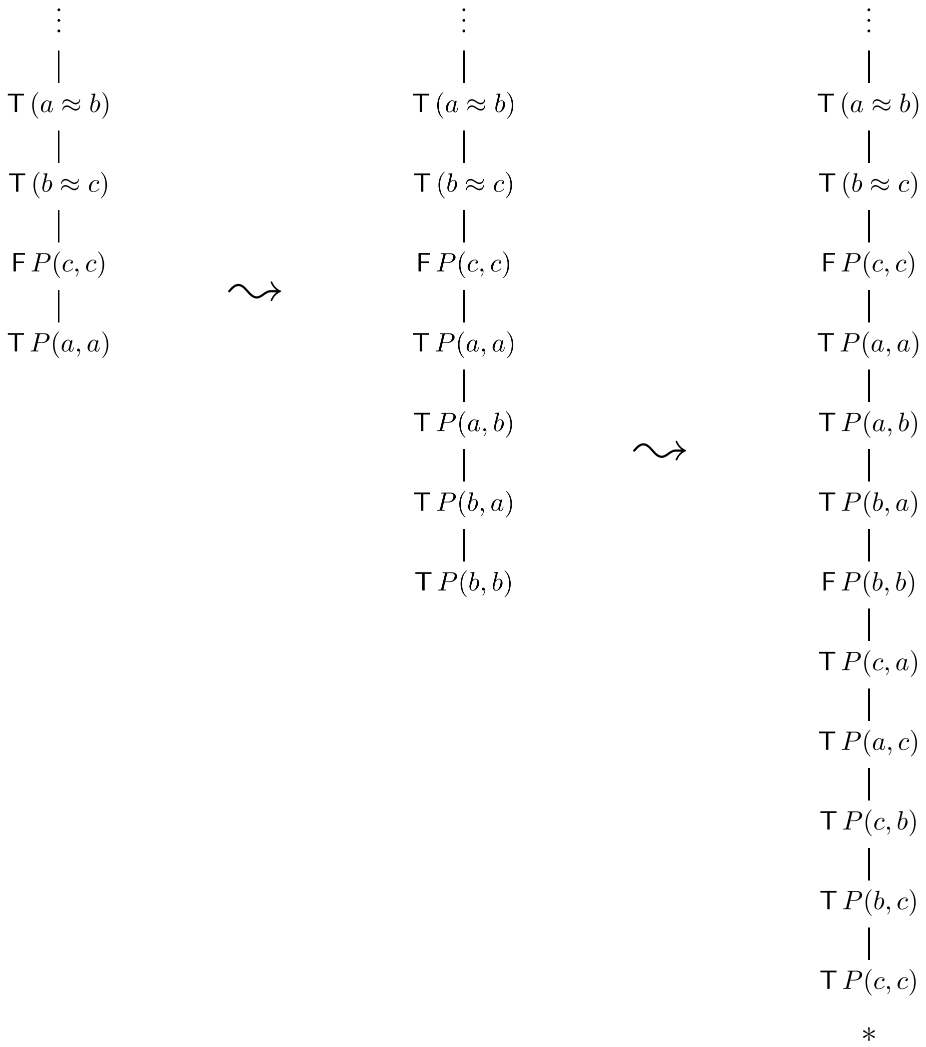


Abbildung 4.4: Abschluß eines Astes nach dem Verfahren von Jeffrey.

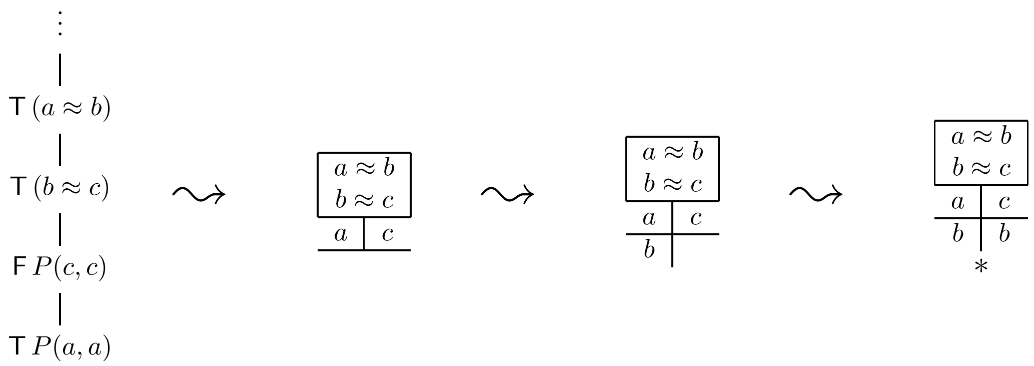


Abbildung 4.5: Abschluß desselben Astes wie in Abbildung 4.4, hier jedoch nach der Methode der Umwandlung in Disjunktionen von Ungleichungen. Dargestellt sind jeweils in einem Rechteck die anwendbaren Gleichungen, darunter die Ungleichung, von der ausgegangen wird, und unter den beiden Termen der Ungleichung die aus ihnen abgeleiteten Terme.

die richtige Reihenfolge der Gleichungsanwendungen auf den Ast bzw. die richtige Reihenfolge der sukzessiven Berechnung der Terme in den Äquivalenzklassen zu finden, außer acht gelassen. Sie sei für die beiden gegenübergestellten Klassen von Verfahren in gleicher Weise durch eine Heuristik gelöst. Diese bewirke, daß n Anwendungen von Gleichungen aus $\mathrm{Gl}(A)$ benötigt werden, um zu zeigen, daß in der durch $\mathrm{Gl}(A)$ gegebenen Gleichungstheorie $t_i = s_i$ für $1 \leq i \leq k$ gilt.

Dann kann der Ast nach einer Methode, die auf der Erweiterung des Astes beruht, etwa der von Jeffrey, mit

$$o(k \cdot n) \text{ und } O(n^k)$$

Erweiterungsschritten[2] geschlossen werden. Der schlechteste Fall tritt dann ein, wenn bei der Ableitung, wie in Abbildung 4.4 zu Beispiel 4.3.8, auch die unnötigen Zwischenschritte auftreten. Der beste Fall wird erreicht, wenn eine optimale Heuristik zur Verfügung steht, die alle diese unnötigen Zwischenschritte vermeidet.

Dieser beste Fall wird bei einem Abschluß des Astes mit Hilfe von Äquivalenzklassen aber immer erreicht, denn dafür werden

$$o(n) \text{ und } O(k \cdot n)$$

Erweiterungsschritte benötigt. Der optimale Fall einer Komplexität von $o(n)$ wird erreicht, wenn die Ungleichungen $t_i \not\approx s_i$ für $1 \leq i \leq k$ identisch sind (wie in Beispiel 4.3.8).

Dies alles gilt allerdings nur unter der Voraussetzung, daß die den Abschluß ermöglichende Substitution σ schon gefunden ist. Das Problem, eine solche Substitution zu bestimmen, wird in den nächsten Abschnitten behandelt.

4.4 Abschluß mit Gleichheit und universellen Formeln

Bevor näher auf die Suche nach einer den Abschluß eines Tableaus erlaubenden Substitution eingegangen wird, soll das bisher beschriebene Verfahren so erweitert werden, daß beim Abschluß die Universalität von Formeln und insbesondere auch von Gleichungen berücksichtigt wird.

Im Gegensatz zu dem in Kapitel 3 eingeführten Verfahren zur Berücksichtigung universeller Formeln, soll nun aber die Definition nicht in der Weise verändert werden, daß eine abschließende Substitution für jeden Ast anzugeben ist. Vielmehr soll auch weiter eine einzige das ganze Tableau abschließende Substitution gesucht werden.

Die Universalität von Gleichungen geht direkt in die Menge $\mathrm{Gl}(A\sigma)$ der Gleichungen eines Astes A ein, deren Definition in der folgenden Weise verändert wird, und die dann mit $\mathrm{Gl}^*(A, \sigma)$ bezeichnet wird:

Definition 4.4.1 (Menge der Gleichungen eines Astes mit universellen Gleichungen)

Seien A der Ast eines Tableaus und σ eine Grundsubstitution. Ferner sei $\mathcal{M}$ eine Methode zum Auffinden universeller Formeln.

Eine Menge der Form

$$\{((\forall x_1)\dots(\forall x_n)(s \approx t))\sigma \;:\; s \approx t \in A,\, \mathcal{M}(A, s \approx t) = \{x_1, \dots, x_n\}\,\}$$

wird mit $\mathrm{Gl}^*(A, \sigma)$ *bezeichnet.*

Eine Menge $\mathrm{Gl}^*(A, \sigma)$ enthält also die Gleichungen des Astes $A\sigma$, wobei allerdings in Gleichungen, die universell sind bezüglich einer Variablen x, die Variable x von der Substitution ausgenommen und allquantifiziert wird, was gerade der Idee der universellen Formeln entspricht.

[2] o bezeichnet die Komplexität im besten, O die Komplexität im schlechtesten Fall.

$\mathrm{Gl}^*(A,\sigma)$ ist nicht wie $\mathrm{Gl}(A\sigma)$ eindeutig definiert. Welche der möglichen Mengen vom Typ $\mathrm{Gl}^*(A,\sigma)$ man erhält, hängt von $\mathcal{M}$ ab, also davon, welche der Gleichungen auf dem Ast A man als universell erkennt. Insbesondere ist $\mathrm{Gl}(A\sigma)$ eine Menge vom Typ $\mathrm{Gl}^*(A,\sigma)$.

Beispiel 4.4.2

Definition 4.4.1 soll wieder am Beispiel des Astes A_{Bsp} verdeutlicht werden. Zugrunde gelegt wird das noch freie Variablen enthaltende Tableau, wie es in Abbildung 4.1 (a) dargestellt ist. Als Methode zum Auffinden universeller Formeln wird diejenige aus Satz 3.1.3 verwendet.

Zunächst ist nun noch die Substitution σ zu wählen, für die $\mathrm{Gl}^*(A_{\mathrm{Bsp}},\sigma)$ angegeben werden soll. Da die Gleichung $\mathsf{T}\,(g(f(x_1)) \approx x_1)$ als universell bezüglich x_1 erkannt wird, ist die Belegung von x_1 unerheblich. Sei etwa $\sigma = \{x_1 \leftarrow a,\, x_2 \leftarrow a\}$. Damit ist

$$\mathrm{Gl}^*(A_{\mathrm{Bsp}},\sigma) = \{b \approx c,\, (\forall x_1)(g(f(x_1)) \approx x_1),\, g(a) \approx f(a)\}.$$

Für alle weiteren Überlegungen sei nun aber, um Unklarheiten zu vermeiden, eine bestimmte Methode $\mathcal{M}$ zum Auffinden universeller Formeln ausgewählt, und damit auch die Menge $\mathrm{Gl}^*(A,\sigma)$ eindeutig definiert.

Nun kann auch die Definition der Äquivalenzklasse eines Terms so erweitert werden, daß die Universalität von Gleichungen berücksichtigt wird.

Definition 4.4.3 (Äquivalenzklasse, basierend auf universellen Gleichungen)

Sei t ein geschlossener Term, A der Ast eines Tableaus und σ eine Grundsubstitution.

Dann bezeichne $[t]^*_{(A,\sigma)}$ die Äquivalenzklasse von t bezüglich der durch die Gleichungen in $\mathrm{Gl}^*(A,\sigma)$ gegebenen Gleichungstheorie, also bezüglich der Relation, die definiert ist durch

$t \sim s$ *gdw.* $t = s$ *gilt in der durch die Gleichungen in* $\mathrm{Gl}^*(A,\sigma)$ *gegebenen Gleichungstheorie.*

Es soll aber nicht nur die Universalität der Gleichungen auf einem Ast A, sondern auch diejenige der Formeln berücksichtigt werden, aus denen die Disjunktionen in $\mathrm{Dis}(A)$ gebildet werden. Dies kann in einfacher Weise dadurch geschehen, daß in einer Disjunktion $D \in \mathrm{Dis}(A)$ auftretende Variablen, bezüglich derer die Formeln, aus denen D entstanden ist, universell sind, durch neue, nicht an anderer Stelle schon auftretende Variablen ersetzt werden:

Definition 4.4.4 (Bildung der Mengen $\mathrm{Dis}^*(A_i)$ aus $\mathrm{Dis}(A_i)$)

Ein Tableau T bestehe aus den Ästen $A_1,\ldots,A_k$. Für $1 \le i \le k$ mögen die Mengen $\mathrm{Dis}(A_i)$ die Form

$$\mathrm{Dis}(A_i) = \{D_{i1},\ldots,D_{im_i}\}$$

haben. $\{y^1_{ij},\ldots,y^{n_{ij}}_{ij}\}$ sei die Menge der Variablen, bezüglich derer die Formeln, aus denen D_{ij} entstanden ist, als universell erkannt werden ($1 \le j \le m_i$).

Dann sind für $1 \le i \le k$ die Mengen $\mathrm{Dis}^*(A_i)$ gegeben durch

$$\mathrm{Dis}^*(A_i) = \{D_{ij}\{y^1_{ij} \leftarrow u^1_{ij},\ldots,y^{n_{ij}}_{ij} \leftarrow u^{n_{ij}}_{ij}\} \;:\; 1 \le j \le m_i\}.$$

Dabei sind die Variablen u^l_{ij} neu, d. h. sie sind paarweise verschieden und treten in T nicht auf.

Man beachte, daß eine Variable, wenn sie in verschiedenen Disjunktionen auftritt, zwar dort durch verschiedene neue Variablen ersetzt werden kann, die verschiedenen Vorkommen einer Variablen innerhalb einer Disjunktion muß aber immer durch die gleiche Variable ersetzt werden.

Beispiel 4.4.5

Es sei möglich, eine Disjunktion $D_i \in \mathrm{Dis}(A_i)$ mit Hilfe der Substitution $\{x \leftarrow a\}$ und eine Disjunktion $D_j \in \mathrm{Dis}(A_j)$ mit Hilfe von $\{x \leftarrow b\}$ zu widerlegen, und die Formeln aus denen die beiden Disjunktionen entstanden sind, seien universell bezüglich x.

Dann ist es nach der Substitution von x durch neue Variablen u_i bzw. u_j möglich, beide Disjunktionen mit Hilfe der einen Substitution $\{u_i \leftarrow a,\, u_j \leftarrow b\}$ zu widerlegen, was, da die Disjunktionen universell bezüglich der ursprünglichen Variablen x sind, korrekt ist.

Der Abschluß eines Tableaus mit universellen Formeln und Gleichheit kann nun formuliert werden:

Definition 4.4.6 (Abschluß eines Tableaus mit Hilfe von Äquivalenzklassen, basierend auf universellen Gleichungen)

Ein Tableau T mit den Ästen $A_1, \ldots, A_k$ ist geschlossen, wenn es eine Grundsubstitution σ gibt und für jeden Ast A_i eine Disjunktion

$$D_i^* = (t_{i1}^* \not\approx s_{i1}^* \vee \ldots \vee t_{in_i}^* \not\approx s_{in_i}^*) \in \mathrm{Dis}^*(A_i),$$

so daß für $1 \leq j \leq n_i$ gilt:

$$[t_{ij}^*\sigma]^*_{(A_i,\sigma)} = [s_{ij}^*\sigma]^*_{(A_i,\sigma)}.$$

Beispiel 4.4.7

In Fortführung von Beispiel 4.4.2 soll wieder das in Abbildung 4.1 (a) dargestellte Tableau mit seinen beiden Ästen A_{Bsp} und A'_{Bsp} betrachtet werden.

Zur Erinnerung:

$$\begin{aligned} \mathrm{Dis}(A_{\mathrm{Bsp}}) &= \{g(g(a)) \not\approx a \vee b \not\approx c\} \\ \mathrm{Dis}(A'_{\mathrm{Bsp}}) &= \{x_2 \not\approx a,\ (g(g(a)) \not\approx a \vee b \not\approx c)\} \end{aligned}$$

Da die Formel $\mathsf{F}\,(x_2 \approx a)$, aus der die Disjunktion $x_2 \not\approx a$ entstanden ist, nicht als universell bezüglich x_2 erkannt wird, und sonst keine freien Variablen auftreten, ist $\mathrm{Dis}^*(A_{\mathrm{Bsp}}) = \mathrm{Dis}(A_{\mathrm{Bsp}})$ und $\mathrm{Dis}^*(A'_{\mathrm{Bsp}}) = \mathrm{Dis}(A'_{\mathrm{Bsp}})$.

Der Abschluß nach Definition 4.4.6 gelingt mit der Substitution

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

denn es gilt:

$$\begin{aligned} [g(g(a))]^*_{(A_{\mathrm{Bsp}},\sigma)} &= \\ [a]^*_{(A_{\mathrm{Bsp}},\sigma)} &= \{a,\ g(g(a)),\ g(f(a)),\ g(g(g(g(a)))),\ g(g(g(f(a)))), \\ &\quad\ g(g(f(g(a)))),\ \ldots\} \\ [b]^*_{(A_{\mathrm{Bsp}},\sigma)} &= \\ [c]^*_{(A_{\mathrm{Bsp}},\sigma)} &= \{b,\ c,\ g(g(b)),\ g(g(c)),\ g(f(b)),\ g(f(c)),\ \ldots\} \\ [a]^*_{(A'_{\mathrm{Bsp}},\sigma)} &= \{a, \ldots\} \end{aligned}$$

Beispiel 4.4.8

Der besondere Vorteil der Verwendung universeller Gleichungen soll auch noch einmal am Beispiel des Assoziativgesetzes der Gruppentheorie gezeigt werden. Vorgegeben sei also der folgende mit der Grenze $q = 1$ für die Anwendung der γ–Regel voll expandierte Ast A_{Bsp2}:

$$\begin{array}{c} \mathsf{F}\,[a \cdot (b \cdot (c \cdot d)) \approx ((a \cdot b) \cdot c) \cdot d] \\ | \\ \mathsf{T}\,(\forall x)(\forall y)(\forall z)(x \cdot y) \cdot z \approx x \cdot (y \cdot z)] \\ | \\ \mathsf{T}\,[(x \cdot y) \cdot z \approx x \cdot (y \cdot z)] \end{array}$$

Wäre die aus der Formel $\gamma = \mathsf{T}\,(\forall x)(\forall y)(\forall z)((x \cdot y) \cdot z \approx x \cdot (y \cdot z))$ abgeleitete Gleichung $G = \mathsf{T}\,((x \cdot y) \cdot z \approx x \cdot (y \cdot z))$ nicht als universell erkannt, könnte A_{Bsp2} nicht geschlossen werden. Ist dies aber wie hier der Fall, so gilt für jede beliebige Substitution σ:

$$\begin{aligned} &\{a \cdot (b \cdot (c \cdot d)),\ \ a \cdot ((b \cdot c) \cdot d),\ \ (a \cdot b) \cdot (c \cdot d),\ \ (a \cdot (b \cdot c)) \cdot d,\ \ ((a \cdot b) \cdot c) \cdot d\} \\ &\quad \subset [(a \cdot b) \cdot (c \cdot d)]^*_{(A_{\mathrm{Bsp2}},\sigma)} \end{aligned}$$

Daraus folgt, da es sich um eine Äquivalenzklasse handelt,

$$[a \cdot (b \cdot (c \cdot d))]^*_{(A_{\mathrm{Bsp2}},\sigma)} = [((a \cdot b) \cdot c) \cdot d]^*_{(A_{\mathrm{Bsp2}},\sigma)}$$

und A_{Bsp2} ist nach Definition 4.4.6 geschlossen.

Es muß noch bewiesen werden, daß das Verfahren mit der neuen Abschlußregel nach Definition 4.4.6 noch korrekt und vollständig ist, was aus dem folgenden Satz folgt:

Satz 4.4.9

Sei T das nach Definition 4.3.6 unter Beachtung der Grenze q für die Anwendung der γ–Regel aufgebaute Tableau für eine Formel X.

Korrektheit: Ist T nach Definition 4.4.6 geschlossen, dann ist die Formel X unerfüllbar.

Vollständigkeit: Ist X unerfüllbar und die Grenze q hinreichend hoch gewählt, dann ist T nach Definition 4.4.6 geschlossen.

Beweis:

Korrektheit:
T sei nach Definition 4.4.6 geschlossen.

Annahme: X ist erfüllbar
Dann ist, da die Erfüllbarkeit bei der Tableauerweiterung erhalten bleibt, auch T erfüllbar (Lemma 4.2.5).

Da T erfüllbar und geschlossen ist nach Definition 4.4.6, hat einer der Äste A_i von T nach Anwendung der Grundsubstitution σ ein Modell $\mathbf{M} = (\mathbf{D}, \mathbf{I})$.

Aus Definition 4.4.1 ergibt sich, daß dann auch die in $\mathrm{Gl}^*(A_i, \sigma)$ enthaltenen Gleichungen in $\mathbf{M}$ Gültigkeit besitzen, woraus wiederum folgt, daß alle Elemente einer Menge $[t]^*_{(A_i,\sigma)}$ durch $\mathbf{I}$ auf das gleiche Element $d \in \mathbf{D}$ abgebildet werden.

Der Ast A_i ist geschlossen. Für $1 \leq i \leq k$ und $1 \leq j \leq n_i$ gilt also

$$[t^*_{ij}\sigma]^*_{(A_i,\sigma)} = [s^*_{ij}\sigma]^*_{(A_i,\sigma)}.$$

Ist die Disjunktion D_i aus einer Ungleichung $\mathsf{F}\,(t_{i1} \approx s_{i1})$ entstanden, folgt

$$(t^*_{i1}\sigma)^{\mathbf{I}} = (s^*_{i1}\sigma)^{\mathbf{I}},$$

und daraus

$$\mathbf{M} \models \mathsf{T}\,(t^*_{i1} \approx s^*_{i1})\sigma.$$

Andererseits ist aber $\mathbf{M}$ als Modell von $A_i\sigma$ auch Modell von $\mathsf{F}\,(t_{i1} \approx s_{i1})\sigma$. Da nach Definition 4.4.4 genau die Variablen durch neue ersetzt wurden, bezüglich derer $\mathsf{F}\,(t_{i1} \approx s_{i1})$ universell ist, gilt auch

$$\mathbf{M} \models \mathsf{F}\,(t^*_{i1} \approx s^*_{i1})\sigma,$$

was zum Widerspruch führt.

In dem Fall, daß die Disjunktion D_i aus Formeln $\mathsf{T}\,P(t_{i1}, \ldots, t_{im_i})$ und $\mathsf{F}\,P(s_{i1}, \ldots, s_{im_i})$ entstanden ist, ist ein ähnlicher Schluß möglich: Zunächst folgert man

$$(P(t^*_{i1}, \ldots, t^*_{in_i})\sigma)^{\mathbf{I}} = (P(s^*_{i1}, \ldots, s^*_{in_i})\sigma)^{\mathbf{I}},$$

und daraus, daß $P(t^*_{i1}, \ldots, t^*_{in_i})\sigma$ und $P(s^*_{i1}, \ldots, s^*_{in_i})\sigma$ in $\mathbf{M}$ den gleichen Wahrheitswert haben müssen. Dies ist aber nicht möglich, denn als Modell von $A_i\sigma$ ist $\mathbf{M}$ auch Modell von $\mathsf{T}\,P(t_{i1}, \ldots, t_{im_i})\sigma$ und $\mathsf{F}\,P(s_{i1}, \ldots, s_{im_i})\sigma$, woraus man, wie oben, gestützt auf Definition 4.4.4,

$$\mathbf{M} \models \mathsf{T}\,P(t^*_{i1}, \ldots, t^*_{im_i})\sigma,\ \mathsf{F}\,P(s^*_{i1}, \ldots, s^*_{im_i})\sigma$$

folgert.

Vollständigkeit:
Diese Richtung des Beweises ist einfacher zu führen, da die neue Abschlußregel eine Abschwächung des Abschlusses ohne universelle Gleichungen (Def. 4.3.4) darstellt.

Sei also die Formel X unerfüllbar und q hinreichend hoch gewählt. Dann ist T nach Definition 4.3.4 geschlossen, was Lemma 4.3.7 zusichert.

Sei σ die nach Definition 4.3.4 existierende Grundsubstitution und $D_1, \ldots, D_k$ die ebenfalls nach Definition 4.3.4 existierenden Disjunktionen aus $\mathrm{Dis}(A_i)$.

Die Substitution ρ kehre die Variablensubstitution nach Definition 4.4.4 um, so daß für $1 \leq i \leq k$ und $1 \leq j \leq m_i$ gilt:

$$t^*_{ij}\rho = t_{ij} \quad \text{und} \quad s^*_{ij}\rho = s_{ij}.$$

Außerdem gilt dann $A_i = A_i\rho$, da die durch ρ belegten Variablen nach Definition 4.4.4 in T nicht vorkommen dürfen. Damit folgt aus

$$[t_{ij}\sigma]_{A_i\sigma} = [s_{ij}\sigma]_{A_i\sigma}$$

die Gleichung

$$[t^*_{ij}\rho\sigma]_{A_i\rho\sigma} = [s^*_{ij}\rho\sigma]_{A_i\rho\sigma}.$$

Dann gilt erst recht

$$[t^*_{ij}\rho\sigma]^*_{(A_i,\sigma\circ\rho)} = [s^*_{ij}\rho\sigma]^*_{(A_i,\sigma\circ\rho)}$$

und also ist T mit der Substitution $\sigma \circ \rho$ nach Definition 4.4.6 geschlossen. ■

Um zu zeigen, daß ein Ast A_i nach Definition 4.4.6 geschlossen ist, muß man für alle Terme t^*, die in $\mathrm{Dis}^*(A_i)$ auftreten, $[t^*\sigma]^*_{(A,\sigma)}$ berechnen, um festzustellen, daß für bestimmte Terme t^* und s^* aus $\mathrm{Dis}^*(A_i)$ die Klassen $[t^*\sigma]^*_{(A_i,\sigma)}$ und $[s^*\sigma]^*_{(A_i,\sigma)}$ gleich sind.

Dies kann dadurch geschehen, daß man sukzessive in $[t^*\sigma]^*_{(A_i,\sigma)}$ und $[s^*\sigma]^*_{(A_i,\sigma)}$ liegende Terme berechnet, bis man auf einen stößt, der in beiden Klassen liegt. Da es sich um Äquivalenzklassen handelt, folgt daraus dann deren Gleichheit.

Wie aus Definition 4.4.6 aber ersichtlich ist, muß nicht nur ein solcher Abschluß eines Astes gesucht werden, sondern auch eine Substitution σ, für die dieser Abschluß Gültigkeit hat. Auf diese Problematik wird im nächsten Abschnitt näher eingegangen.

4.5 Breitensuche nach den Substitutionen

Um eine Substitution zu finden, die den Abschluß eines Astes ermöglicht, könnte man alle möglichen Substitutionen der Reihe nach ausprobieren. Dies wäre aber viel zu ineffizient.

Betrachtet man zunächst nur den Versuch, eine einzelne Ungleichung $t \not\approx s$ einer Disjunktion aus $\mathrm{Dis}^*(A)$ zu widerlegen, bietet sich folgende sehr viel bessere Methode an: Man betrachtet nicht nur Grundsubstitutionen, sondern versucht möglichst allgemeine Substitutionen anzuwenden, die es erlauben, einen Term abzuleiten.

Es bildet sich dann ein Suchbaum nach einer den Abschluß zulassenden Substitution aus. An seiner Wurzel steht die allgemeinste, die leere Substitution *id*. Man beginnt damit, diejenigen Terme zu bestimmen, die sich aus t nach Anwendung von *id* ableiten lassen.

Tritt dann der Fall auf, daß ein Term r' aus einem schon berechneten Term r nur ableitbar ist, wenn beispielsweise eine Substitution $\{x \leftarrow a\}$ ausgeführt wird, verzweigt der Suchbaum. In einem Pfad wird weiter die leere Substitution *id* betrachtet, in dem anderen die Substitution $\{x \leftarrow a\}$.

Beispiel 4.5.1

Als Beispiel soll wieder die Ungleichung $g(g(a)) \not\approx a$ aus $\mathrm{Dis}^*(A_{\mathrm{Bsp}})$ dienen. Mit Hilfe der Gleichungen

$$b \approx c,\ (\forall x_1)(g(f(x_1)) \approx x_1),\ g(x_2) \approx f(x_2)$$

soll also a aus $g(g(a))$ abgeleitet werden, und eine Belegung der Variablen x_1 und x_2 soll bestimmt werden, die dies erlaubt.

Der entstehende Suchbaum ist in Abbildung 4.6 dargestellt. Die kritische Verzweigung ist die Stelle, an der zwischen der günstigen Belegung $x_2 \leftarrow a$ und der ungünstigen Belegung $x_2 \leftarrow f(a)$, die den Abschluß verhindert, gewählt werden muß.

Jeder der Pfade dieses Suchbaumes ist mit einer bestimmten Substitution σ assoziiert — und damit auch allen Belegungen die spezieller sind als σ. Gelingt es in einem Pfad, die Ungleichung zu widerlegen, so ist die mit ihm assoziierte Substitution σ eine allgemeinste, die diesen Widerspruch zuläßt.

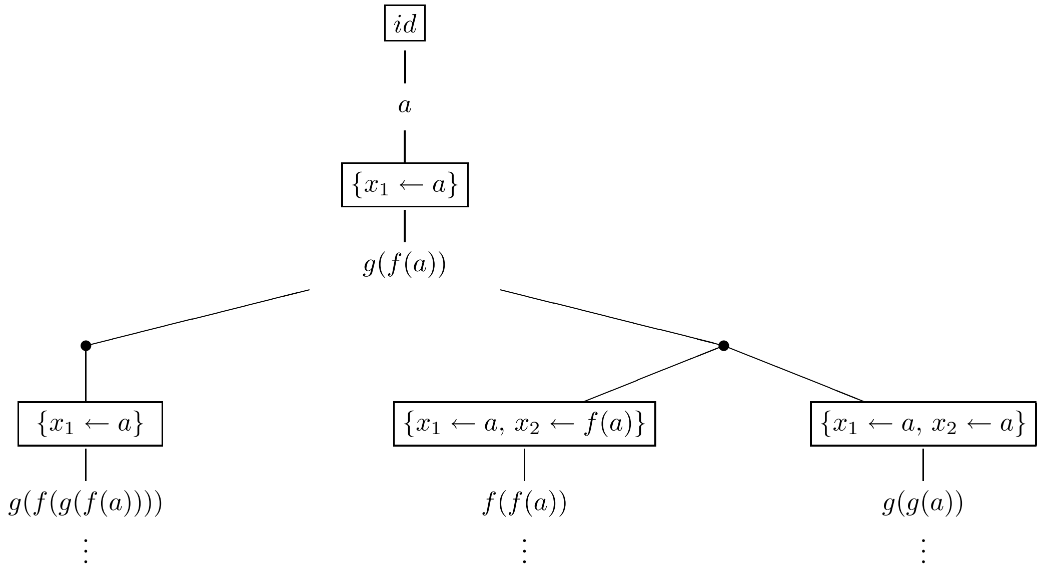


Abbildung 4.6: Suchbaum für die Suche nach einer günstigen Substitution (Beispiel 4.5.1).

Um so den Ast eines Tableaus abzuschließen, genügt es aber nicht, Substitutionen zu finden, nach deren Anwendung die einzelnen Ungleichungen einer Disjunktion widerlegt werden können. Darüber hinaus müssen alle diese Substitutionen auch noch kompatibel sein, so daß ihre Vereinigung es erlaubt, die gesamte Disjunktion zugleich zu widerlegen.

Man muß also gegebenenfalls mehrere abschließende Substitutionen für eine Ungleichung bestimmen, bis man auf eine stößt, die es erlaubt, auch alle anderen Ungleichungen der Disjunktion abzuschließen.

Dieses Verfahren ließe sich in PROLOG relativ einfach implementieren. Man könnte die freien Variablen des Tableaus durch PROLOG–Variablen modellieren, und das Auffinden einer alle Ungleichungen einer Disjunktion abschließenden Substitution dem internen Backtracking des PROLOG überlassen.

Genau so wird bei der Version von $_3T^AP$ ohne Gleichheitsbehandlung vorgegangen, um eine Substitution zu finden, die den Abschluß aller Äste eines Tableaus zugleich gestatten. Zu diesem Zweck wurde das Verfahren auch in der schließlich implementierten Version mit Gleichheitsbehandlung beibehalten (auf die damit verbundene Problematik wird genauer in Abschnitt 6.3 eingegangen).

Um eine Substitution σ zu finden, die den Abschluß eines einzelnen Astes gestattet, wurde es aber nicht übernommen, denn das Verfahren hat einen entscheidenden Nachteil: es modelliert eine Tiefensuche. Variablenbelegungen werden vorgenommen und erst, wenn sie „erkennbar“ zu nichts führen, wird eine andere Variablenbelegung untersucht. Um zu Erkennen, daß eine Substitution ungeeignet ist, gibt es aber kein allgemeines Verfahren, das Problem ist sogar unentscheidbar. Es können nur Heuristiken eingesetzt werden, um einige Spezialfälle abzudecken. In allen anderen Fällen müßte die Suche durch eine feste Schranke begrenzt werden, die nur schwer zu wählen ist. Ist sie zu groß, sind die Äste des Suchbaumes sehr tief. Es dauert dann viel zu lange bis das Backtracking zum Zuge kommt. Wird sie zu klein gewählt, kann unter Umständen ein möglicher Abschluß nicht mehr gefunden werden.

Diese Problematik ist wohl der Hauptgrund dafür, daß Fittings Beweiser so ineffizient ist. Denn auch dort wird eine Tiefensuche nach einer den Abschluß zulassenden Substitution modelliert.

Bei einer Breitensuche nach den abschließenden Substitutionen dagegen werden die meisten der genannten Probleme vermieden. Die Suche wird in allen Pfaden des Suchbaumes zugleich vorangetrieben. Zwar kann Tiefensuche in einzelnen Fällen schneller sein, nämlich dann, wenn gleich zu Anfang im richtigen Pfad gesucht wird. Das kann aber nur dann in der Mehrzahl der Fälle geschehen, wenn der richtige Pfad durch gute Heuristiken ausgewählt wird.

Existieren solche Heuristiken, kann man sie aber auch bei der Breitensuche verwenden, um die Suche in vielversprechenden Pfaden besonders voranzutreiben.

Wenn die Heuristik auch nur an einem einzigen Verzweigungspunkt des Suchbaumes eine falsche Entscheidung trifft, was häufig vorkommen wird, verliert man sich bei der Tiefensuche in dem falsch ausgewählten Pfad, während man durch Breitensuche auch in einem solchen Fall noch eine Lösung finden kann.

Auch ist die Breitensuche sehr viel mächtiger. Durch geschickte Heuristiken kann praktisch jedes andere Suchverfahren simuliert werden. Verbraucht die Breitensuche zuviel Speicherplatz, kann auch dies berücksichtigt werden, indem zunächst nur bestimmte Äste des Suchbaumes vorangetrieben werden.

4.6 Anpassung der Äquivalenzklassen an die Breitensuche

In diesem Abschnitt ist ein Verfahren konzipiert, nach dem die im letzten Abschnitt vorgeschlagene Breitensuche nach Substitutionen, die den Abschluß eines Astes zulassen, ausgeführt werden kann.

Eine Möglichkeit wäre, tatsächlich eine Breitensuche in den im letzen Abschnitt beschriebenen Suchbäumen für die Aquivalenzklassen $[s\sigma]^*_{(A,\sigma)}$ durchzuführen. Dabei müßten aber alle diese Suchbäume zugleich betrachtet werden, was einen enormen Verwaltungs–Overhead bedeuten würde, da das Backtracking des PROLOG nicht mehr verwendet werden kann.

Besser ist ein Ansatz, der die folgende Menge verwendet:

Definition 4.6.1 (Klasse $\langle t\rangle_A$ eines Terms)

Für einen Term t und einen Ast A sei

$$\langle t\rangle_A := \{s_\tau \ : \ \textit{es gibt Grundsubstitutionen } \tau\textit{, so daß } s \in [t\tau]^*_{(A,\tau)}\}$$

$\langle t\rangle_A$ ist selbst keine Äquivalenzklasse mehr, umfaßt aber in gewissem Sinne alle Äquivalenzklassen $[t\sigma]^*_{(A,\sigma)}$, ohne dabei von einer bestimmten Substitution σ abzuhängen.

Nun kann der Abschluß des Tableaus definiert werden, ohne daß die Existenz einer bestimmten Substitution, die den Abschluß zuläßt, und über deren Bestimmung nichts ausgesagt wird, vorausgesetzt würde.

Definition 4.6.2 (Abschluß mit Hilfe von $\langle t\rangle_A$)

Ein Tableau T mit den Ästen $A_1, \ldots, A_k$ ist geschlossen, wenn es für jeden Ast A_i eine Disjunktion

$$D_i^* = (t_{i1}^* \not\approx s_{i1}^* \vee \ldots \vee t_{in_i}^* \not\approx s_{in_i}^*) \in \mathrm{Dis}^*(A_i)$$

gibt, so daß für $1 \leq j \leq n_i$ Elemente

$$r_{\sigma_{ij}}^{ij} \in \langle t_{ij}^*\rangle_{A_i} \cap \langle s_{ij}^*\rangle_{A_i}$$

existieren, die alle dieselbe Substitution $\sigma = \sigma_{ij}$ $(1 \leq i \leq k, 1 \leq j \leq n_i$ besitzen.

Beispiel 4.6.3

Als Beispiel soll wieder der rechte Ast A_{Bsp} des in Abbildung 4.1 (a) dargestellten Tableaus dienen.

Elemente von $\langle g(g(a))\rangle_{A_{\mathrm{Bsp}}}$ sind beispielsweise:

- $g(g(g(f(a))))_{\{x_1 \leftarrow a, x_2 \leftarrow a\}}$
- $g(g(g(f(a))))_{\{x_1 \leftarrow a, x_2 \leftarrow g(a)\}}$
- $f(g(a))_{\{x_1 \leftarrow a, x_2 \leftarrow g(a)\}}$
- $g(f(a))_{\{x_1 \leftarrow a, x_2 \leftarrow a\}}$
- $f(g(g(f(a))))_{\{x_1 \leftarrow a, x_2 \leftarrow g(g(f(a)))\}}$
- $a_{\{x_1 \leftarrow a, x_2 \leftarrow a\}}$

Wie die Variable x_1 belegt wird, ist in all diesen Fällen unwichtig, da die einzige Formel, in der x_1 vorkommt, die Gleichung $\mathsf{T}\,(g(f(x_1)) \approx x_1)$, universell ist bezüglich x_1.

$\langle g(g(a)) \rangle_{A_{\text{Bsp}}}$ und $\langle a \rangle_{A_{\text{Bsp}}}$ haben ein gemeinsames Element, nämlich

$$r^{1,1}_{\sigma_{1,1}} = a_{\{x_1 \leftarrow a, x_2 \leftarrow a\}}.$$

Weiter liegt

$$r^{1,2}_{\sigma_{1,2}} = b_{\{x_1 \leftarrow a, x_2 \leftarrow a\}}$$

im Schnitt von $\langle b \rangle_{A_{\text{Bsp}}}$ und $\langle c \rangle_{A_{\text{Bsp}}}$.

Mit der Substitution $\sigma = \sigma_{1,1} = \sigma 2,2 = \{x_1 \leftarrow a,\, x_2 \leftarrow a\}$ ist die Disjunktion

$$(g(g(a)) \not\approx a \vee b \not\approx c) \in \mathrm{Dis}^*(A_{\text{Bsp}})$$

also widerlegt. Das gleiche gilt für die Disjunktion $x_2 \not\approx a$ aus $\mathrm{Dis}^*(A'_{\text{Bsp}})$ und das Tableau ist nach Definition 4.6.2 mit der Substitution σ geschlossen.

Der folgende Satz stellt Korrektheit und Vollständigkeit des Abschlusses nach Definition 4.6.2 sicher.

Satz 4.6.4

Sei T das nach Definition 4.3.6 unter Beachtung der Grenze q für die Anwendung der γ–Regel aufgebaute Tableau für eine Formel X.

Korrektheit: Ist T nach Definition 4.6.2 geschlossen, dann ist die Formel X unerfüllbar.

Vollständigkeit: Ist X unerfüllbar und die Grenze q hinreichend hoch gewählt, dann ist T nach Definition 4.6.2 geschlossen.

Beweis:

Der Abschluß nach Definition 4.6.2 wird auf den nach Definition 4.4.6 zurückgeführt.

Korrektheit:
Sei T nach Definition 4.6.2 geschlossen, dann gibt es für jeden seiner Äste A_i $(1 \leq i \leq k)$ eine Disjunktion

$$D^*_i = (t^*_{i1} \not\approx s^*_{i1} \vee \ldots \vee t^*_{in_i} \not\approx s^*_{in_i}) \in \mathrm{Dis}^*(A_i),$$

so daß für $1 \leq j \leq n_i$ Elemente

$$r^{ij}_{\sigma_{ij}} \in \langle t^*_{ij} \rangle_{A_i} \cap \langle s^*_{ij} \rangle_{A_i}$$

existieren mit der gleichen Substitution $\sigma_{ij} = \sigma$. Mit Definition 4.6.1 folgert man sofort

$$[t^*_{ij}\sigma]^*_{(A_i,\sigma)} = [s^*_{ij}\sigma]^*_{(A_i,\sigma)}.$$

T ist also auch nach Definition 4.4.6 geschlossen, woraus die Unerfüllbarkeit von X folgt (Satz 4.4.9).

Vollständigkeit:
Sei X unerfüllbar und q hinreichend hoch gewählt. Dann ist T nach Definition 4.4.6 geschlossen. Es gibt also eine Substitution σ und Disjunktionen $D^*_i \in \mathrm{Dis}^*(A_i)$ $(1 \leq i \leq k)$, so daß

$$[t^*_{ij}\sigma]^*_{(A_i,\sigma)} = [s^*_{ij}\sigma]^*_{(A_i,\sigma)}.$$

Da es sich um nichtleere Äquivalenzklassen handelt, folgert man

$$(t^*_{ij})_\sigma \in \langle t^*_{ij} \rangle_{A_i} \cap \langle s^*_{ij} \rangle_{A_i}.$$

Also ist T nach Definition 4.6.2 geschlossen. ■

Nun ist das Problem, die richtige Substitution zu finden, aber immer noch nicht ganz gelöst. Es ist nur in die Berechnung der Menge $\langle t \rangle_A$ verlagert. Wie diese Berechnung aber erfolgt, wird nicht vorgegeben. Beliebige Suchverfahren nach Termen in $\langle t \rangle_A$ können implementiert werden.

t_σ	$t'_{\sigma'}$	τ
a_{id}	a_{id}	id
$a_{\{x \leftarrow y\}}$	$a_{\{x \leftarrow b, y \leftarrow b\}}$	$\{y \leftarrow b\}$
x_{id}	$a_{\{x \leftarrow a\}}$	$\{x \leftarrow a\}$
$f(x)_{\{y \leftarrow a\}}$	$f(a)_{\{x \leftarrow a, y \leftarrow a\}}$	$\{x \leftarrow a\}$

Tabelle 4.1: Beispiele für die Subsumtion nach Definition 4.7.1.

4.7 Berechnung der Menge $\langle t \rangle_A$

Im folgenden ist das Verfahren zur Suche nach dem Abschluß eines Tableaus mit Hilfe von Gleichheit beschrieben, wie es schließlich implementiert wurde. Es verwendet die im letzten Abschnitt definierten Mengen $\langle t \rangle_A$.

Die Menge $\langle t \rangle_A$ tatsächlich zu berechnen, wäre zu aufwendig, denn von den Substitutionen τ, für die ein Term s_τ in $\langle t \rangle_A$ liegt, wird verlangt, daß sie Grundsubstitutionen sind.

Statt also alle in $\langle t \rangle_A$ liegenden Terme sukzessive zu berechnen, wird eine Folge von endlichen Mengen $\langle t \rangle_A^n$ mit $\langle t \rangle_A^0 = \{t_{id}\}$ bestimmt, die $\langle t \rangle_A$ annähert.

Dabei werden jedoch solche Terme t_σ weggelassen, die nur Abschlüsse ermöglichen, die auch ein anderer Term $t'_{\sigma'} \in \langle t \rangle_A$ ermöglicht, der aber t_σ subsumiert, also in gewissem Sinne allgemeiner ist.

Definition 4.7.1 (Subsumtion von Termen)

Seien t und t' Terme und σ und σ' Substitutionen.

t_σ subsumiert $t'_{\sigma'}$, wenn es eine kanonische Substitution τ gibt, so daß

$$t\tau = t'$$

gilt, und σ' eine gemeinsame Spezialisierung von σ und τ ist.

In Zeichen: $t_\sigma \gg t'_{\sigma'}$.

Beispiel 4.7.2

In Tabelle 4.1 sind einige Beispiele für die Subsumtion von Termen nach Definition 4.7.1 angegeben.

Nicht subsumiert wird dagegen beispielsweise

- $a_{x \leftarrow f(b)}$ durch $a_{x \leftarrow f(y)}$ oder
- x_{id} durch y_{id}.

Eine wichtige Eigenschaft der Subsumtion nach Definition 4.7.1 ist ihre Transitivität.

Lemma 4.7.3

Es gelte $t_\sigma \gg t'_{\sigma'}$ und $t'_{\sigma'} \gg t''_{\sigma''}$. Dann gilt auch $t_\sigma \gg t''_{\sigma''}$.

Beweis:

Es gelte $t_\sigma \gg t'_{\sigma'}$ und $t'_{\sigma'} \gg t''_{\sigma''}$. Dann folgt aus Definition 4.7.1, daß es Substitutionen τ und τ' gibt, so daß

$$t\tau = t' \text{ und } t'\tau' = t''$$

gilt und $\sigma \succ \sigma' \succ \sigma''$, $\tau \succ \sigma'$ und $\tau' \succ \sigma''$.

Sei nun $\tau'' = (\tau' \circ \tau)$. Damit ist

$$t\tau'' = (t\tau)\tau' = t'\tau' = t''.$$

Außerdem ist $\sigma \succ \sigma' \succ \sigma''$, und daher auch $\sigma \succ \sigma''$.

Schließlich gilt $\tau'' \succ \sigma''$, denn es gibt Substitutionen μ und ν, so daß

$$\sigma'' = \mu \circ \tau = \nu \circ \tau',$$

woraus

$$\mu \circ \tau'' = \nu \circ \tau' \circ \tau = \mu \circ \tau \circ \tau = \mu \circ \tau = \sigma''$$

folgt. Dabei muß man beachten, daß τ nach Definition 4.7.1 ebenfalls kanonisch ist, und daher $\tau \circ \tau = \tau$ gilt. ∎

Die Vereinfachung, auf Terme zu verzichten, die von anderen subsumiert werden, wird dadurch ermöglicht, daß bei den Elementen s_τ einer Menge $\langle t\rangle_A^n$ — im Gegensatz zu $\langle t\rangle_A$ — die Substitutionen τ keine Grundsubstitutionen zu sein brauchen.

Um $\langle t\rangle_A^n$ zu $\langle t\rangle_A^{n+1}$ zu erweitern, wird aus $\langle t\rangle_A^n$ durch eine Heuristik $\mathcal{H}$ ein Term $s_\sigma = \mathcal{H}(\langle t\rangle_A^n) \in \langle t\rangle_A^n$ ausgewählt. Es werden dann die Terme hinzugefügt (sofern sie nicht durch andere Terme subsumiert werden), die aus s_σ in einem Schritt abgeleitet werden können.

Die folgende Definition faßt diese Beschreibung noch einmal zusammen:

Definition 4.7.4 (Die Mengen $\langle t\rangle_A^n$)

Die Folge der Mengen $\langle t\rangle_A^n$ $(n \geq 0)$ ist definiert durch

- $\langle t\rangle_A^0 = \{t_{id}\}$.
- *Die Menge Θ_n sei gegeben durch*

  $$\Theta_n = \{s_\tau \;:\; s_\tau \text{ ist in einem Schritt aus } \mathcal{H}(\langle t\rangle_A^n) \text{ ableitbar}\}$$

  Dabei ist ein Element s_τ aus $t'_\sigma = \mathcal{H}(\langle t\rangle_A^n)$ in einem Schritt ableitbar, wenn

  1. *s aus $t'\tau$ mit Hilfe einer Gleichung $G \in \mathrm{Gl}^*(A,\tau)$ in einem Schritt ableitbar ist,*
  2. *τ eine Spezialisierung von σ ist, und*
  3. *τ eine allgemeinste Substitution ist, die die Bedingungen 1 und 2 erfüllt.*

  $\langle t\rangle_A^n$ wird in folgender Weise aus Θ_n bestimmt:

  $$\begin{array}{rcl} \langle t\rangle_A^{n+1} & = & (\langle t\rangle_A^n \cup \Theta_n) \setminus \\ & & \{s_\tau \;:\; \textit{es gibt } s'_{\tau'} \in (\langle t\rangle_A^n \cup \Theta_n) \textit{ mit} \\ & & \quad s'_{\tau'} \textit{ subsumiert } s_\tau\textit{, und es gilt nicht, daß umge-} \\ & & \quad \textit{kehrt } s'_{\tau'} \textit{ von } s \textit{ subsumiert wird, und } s_\tau \textit{ lexiko-} \\ & & \quad \textit{graphisch geordnet kleiner ist als } s'_{\tau'}\ \}. \end{array}$$

Im letzten Teil der Definition ist festgelegt, daß von mehreren Termen, die sich gegenseitig subsumieren, genau einer ausgewählt und zu der Menge hinzugefügt wird.

Die Menge Θ_n kann unendlich sein, nämlich dann, wenn mit Hilfe einer Gleichung G, die universell ist bezüglich einer Variablen x, unendlich viele Elemente s_τ in einem Schritt abgeleitet werden können (Beispiel 4.7.5). Die Menge $\langle t\rangle_A^{n+1}$ ist dennoch endlich, da stets nur endlich viele Elemente in Θ_n sich nicht gegenseitig subsumieren. Will man $\langle t\rangle_A^{n+1}$ algorithmisch bestimmen, ist es darum günstiger, statt Θ_n eine Menge Θ'_n zu berechnen, in die von unendlich vielen Elementen $(s(r))_\tau$, die sich nur dadurch unterscheiden, daß beliebige Terme r für x eingesetzt sind, nur das eine Element $(s(u))_\tau$ in Θ'_n aufgenommen wird. Dabei ist u eine neue, bisher noch nicht aufgetretene Variable. $\langle t\rangle_A^{n+1}$ kann dann in gleicher Weise aus Θ'_n wie aus Θ_n bestimmt werden.

n	$\langle a \rangle^n_{A_{\text{Bsp}}}$	$\mathcal{H}(\langle a \rangle^n_{A_{\text{Bsp}}})$	Θ_n
0	a_{id}	a_{id}	$g(f(a))_{id}$
1	a_{id} $g(f(a))_{id}$	$g(f(a))_{id}$	a_{id} $g(f(g(f(a))))_{id}$ $g(g(f(f(a))))_{id}$ $g(g(a))_{\{x_2 \leftarrow a\}}$ $f(f(a))_{\{x_2 \leftarrow f(a)\}}$
2	a_{id} $g(f(a))_{id}$ $g(f(g(f(a))))_{id}$ $g(g(f(f(a))))_{id}$ $g(g(a))_{\{x_2 \leftarrow f(a)\}}$ $f(f(a))_{\{x_2 \leftarrow a\}}$		

Tabelle 4.2: Berechnung von $\langle a \rangle^n_{A_{\text{Bsp}}}$ $(n = 0, 1, 2)$.

Beispiel 4.7.5

Unendlich viele Elemente s_τ können beispielsweise aus a_{id} mit Hilfe der Gleichung $(\forall x)(a \approx f(x))$ abgeleitet werden. In Θ'_n wird in diesem Fall aber nur das Element $f(u)_{id}$ aufgenommen werden.

Wichtig ist festzuhalten, daß sich die Menge Θ'_n in relativ einfacher Weise algorithmisch bestimmen läßt und damit $\langle t \rangle^{n+1}_A$ aus $\langle t \rangle^n_A$ zu berechnen ist — nämlich dadurch, daß man alle Möglichkeiten betrachtet, eine der Gleichungen aus $\mathrm{Gl}^*(A, \sigma)$ auf $\mathcal{H}(\langle t \rangle^n_A)$ anzuwenden, und die dafür notwendigen Substitutionen mit Hilfe eines Unifikationsalgorithmus bestimmt.

Beispiel 4.7.6

Betrachtet werden soll wieder das Beispiel der Ungleichung $g(g(a)) \not\approx a$ aus $\mathrm{Dis}^*(A_{\text{Bsp}})$ (Abbildung 4.1 (a)).

Zur Erinnerung:

$$\mathrm{Gl}^*(A_{\text{Bsp}}, \sigma) = \{b \approx c,\ (\forall x_1)(g(f(x_1)) \approx x_1),\ g(x_2)\sigma \approx f(x_2)\sigma\}.$$

Es werde eine Auswahlheuristik $\mathcal{H}$ mit der Länge eines Terms als einzigem Kriterium benutzt. Ausgewählt wird jeweils der kürzeste nicht schon zuvor ausgewählte Term.

In Tabelle 4.2 ist die Berechnung der Menge $\langle a \rangle^n_{A_{\text{Bsp}}}$ und in Tabelle 4.3 die der Menge $\langle g(g(a)) \rangle^n_{A_{\text{Bsp}}}$ jeweils für $n = 0, 1, 2$ dargestellt.

Jetzt kann der Abschluß des Astes festgestellt werden, beispielsweise mit den Elementen $g(g(a))_{\{x_2 \leftarrow a\}} \in \langle a \rangle^2_{A_{\text{Bsp}}}$ und $g(g(a))_{id} \in \langle g(g(a)) \rangle^0_{A_{\text{Bsp}}}$, deren Substitutionen sich zu der die Ungleichung $a \not\approx g(g(a))$ widerlegenden und den Ast abschließenden Substitution $\{x_2 \leftarrow a\}$ vereinigen lassen.

Damit das Verfahren vollständig ist, mit Hilfe von $\langle t \rangle^n_A$ also tatsächlich alle Abschlüsse gefunden werden können, muß die Auswahlheuristik $\mathcal{H}$ in dem Sinne fair sein, daß durch sie jeder Term t oder ein anderer Term, der t subsumiert, früher oder später ausgewählt wird:

Definition 4.7.7 (Fairneß einer Auswahlheuristik)

Eine Auswahlheuristik $\mathcal{H}$ ist fair, wenn für jeden Term t, jedes $n \geq 0$ und jedes Element s_σ aus der Menge $\langle t \rangle^n_A$ ein $m \geq 0$ existiert, so daß

$$\mathcal{H}(\langle t \rangle^m_A) = s'_{\sigma'}$$

und $s'_{\sigma'}$ subsumiert s_σ.

Eine faire Auswahlheuristik kann beispielsweise dadurch festgelegt werden, daß man auf den Termen eine geeignete Gewichtsfunktion definiert und dann jeweils den leichtesten, nicht schon zuvor gewählten Term auswählt.

n	$\langle g(g(a))\rangle^n_{A_{\mathrm{Bsp}}}$	$\mathcal{H}(\langle g(g(a))\rangle^n_{A_{\mathrm{Bsp}}})$	Θ_n
0	$g(g(a))_{id}$	$g(g(a))_{id}$	$g(f(g(g(a))))_{id}$ $g(g(f(g(a))))_{id}$ $g(g(g(f(a))))_{id}$ $f(g(a))_{\{x_2\leftarrow g(a)\}}$ $g(f(a))_{\{x_2\leftarrow a\}}$
1	$g(g(a))_{id}$ $g(f(g(g(a))))_{id}$ $g(g(f(g(a))))_{id}$ $g(g(g(f(a))))_{id}$ $f(g(a))_{\{x_2\leftarrow g(a)\}}$ $g(f(a))_{\{x_2\leftarrow a\}}$	$f(g(a))_{\{x_2\leftarrow g(a)\}}$	$g(f(f(g(a))))_{\{x_2\leftarrow g(a)\}}$ $f(g(f(g(a))))_{\{x_2\leftarrow g(a)\}}$ $f(g(g(f(a))))_{\{x_2\leftarrow g(a)\}}$ $g(g(a))_{\{x_2\leftarrow g(a)\}}$
2	$g(g(a))_{id}$ $g(f(g(g(a))))_{id}$ $g(g(f(g(a))))_{id}$ $g(g(g(f(a))))_{id}$ $f(g(a))_{\{x_2\leftarrow g(a)\}}$ $g(f(a))_{\{x_2\leftarrow a\}}$ $g(f(f(g(a))))_{\{x_2\leftarrow g(a)\}}$ $f(g(f(g(a))))_{\{x_2\leftarrow g(a)\}}$ $f(g(g(f(a))))_{\{x_2\leftarrow g(a)\}}$		

Tabelle 4.3: Berechnung von $\langle g(g(a))\rangle^n_{A_{\mathrm{Bsp}}}$ $(n = 0, 1, 2)$.

Satz 4.7.8

Sei G eine Gewichtsfunktion auf **Term**, *für die*

1. *für jedes $w \geq 0$ die Menge*

$$\{t \in \mathbf{Term} \,:\, G(t) \leq w\},$$

endlich ist, und die

2. *monoton ist bezüglich der Subsumtion, d. h.*

wenn $t \gg t'$, dann gilt $G(t) \leq G(t')$.

Dann ist die mit Hilfe von G gebildete Auswahlheuristik, die jeweils einen

1. *leichtesten,*
2. *nicht schon einmal ausgewählten*

Term auswählt fair im Sinne von Definition 4.7.7.

Beweis:

Sei $s_\sigma \in \langle t\rangle^n_A$ beliebig. Zunächst ist festzuhalten, daß es für alle $m \geq n$ ein $s'_{\sigma'} \in \langle t\rangle^m_A$ gibt, das s_σ subsumiert, denn nach Definition 4.7.4 kann ein Element von $\langle t\rangle^{n+k}_A$ nur dann aus $\langle t\rangle^{n+k+1}_A$ herausfallen, wenn es von einem Element von $\langle t\rangle^{n+k+1}_A$ subsumiert wird.

Für alle diese $s'_{\sigma'} \in \langle t\rangle^m_A$ gilt $G(s') \leq G(s)$.

Würde nicht s_σ selbst schon für ein $m < n$ ausgewählt, und würde nicht schließlich eines der $s'_{\sigma'}$ ausgewählt, dann müßte für jedes $m \geq 0$ ein anderer Term aus $\langle t\rangle^{n+m}_A$ ausgewählt werden, dessen Gewicht ebenfalls kleiner oder gleich $G(s)$ ist. Dies ist aber nicht möglich, da jeder Term nur einmal ausgewählt werden kann, und die Zahl der Terme, deren Gewicht kleiner oder gleich $G(s)$ ist, als endlich vorausgesetzt wurde. ∎

Beispiel 4.7.9

Ein einfaches Beispiel für eine Satz 4.7.8 entsprechende Gewichtsfunktion ist die syntaktische Länge der Terme, womit eine faire Auswahlheuristik zur Verfügung steht.

Die folgende Definition beschreibt den Abschluß eines Tableaus mit Hilfe der Mengen $\langle A \rangle_t^n$.

Definition 4.7.10 (Abschluß eines Tableaus mit Hilfe von $\langle A \rangle_t^n$)

Ein Tableau T mit den Ästen $A_1, \ldots, A_k$ ist geschlossen, wenn folgendes gilt:

1. *Für jeden Ast A_i $(1 \leq i \leq k)$ gibt es eine Disjunktion*

$$D_i^* = (t_{i1}^* \not\approx s_{i1}^* \vee \ldots \vee t_{in_i}^* \not\approx s_{in_i}^*) \in \mathrm{Dis}^*(A_i),$$

so daß es für $1 \leq j \leq n_i$ Elemente $r_{\rho_{ij}}^{ij}$ und $\overline{r}_{\overline{\rho}_{ij}}^{ij}$ und Indizes $l_{ij}, \overline{l}_{ij} \geq 0$ gibt mit

$$r_{\rho_{ij}}^{ij} \in \langle t_{ij}^* \rangle_{A_i}^{l_{ij}} \quad \textit{und} \quad \overline{r}_{\overline{\rho}_{ij}}^{ij} \in \langle s_{ij}^* \rangle_{A_i}^{\overline{l}_{ij}},$$

und r^{ij} und $\overline{r}^{ij}$ unifizierbar sind mit einem MGU $\overline{\overline{\rho}}_{ij}$.

2. *Die Menge der Substitutionen*

$$\{\rho_{ij}, \overline{\rho}_{ij}, \overline{\overline{\rho}}_{ij} \; : \; 1 \leq i \leq k, \, 1 \leq j \leq n_i\}$$

eine kanonische Vereinigung σ besitzt.

Die in Definition 4.7.10 aufgestellte Forderung, daß die Substitution σ kanonisch sein soll, behindert die Suche nach Abschlüssen nicht. In der Praxis werden alle beteiligten Substitutionen mit Hilfe von Unifikationsalgorithmen bestimmt, die ohnehin nur kanonische Substitutionen liefern.

Zum Beweis der Korrektheit und der Vollständigkeit dieses Abschlusses werden die beiden folgenden Lemmata benötigt, die den Zusammenhang verdeutlichen, der zwischen $\langle t \rangle_A$ und der Folge der Mengen $\langle t \rangle_A^n$ besteht.

Lemma 4.7.11

Für ein $n \geq 0$ sei

$$r_\sigma \in \langle t \rangle_A^n.$$

Dann gilt für jede Grundsubstitution τ, die eine Spezialisierung von σ ist,

$$(r\tau)_\tau \in \langle t \rangle_A.$$

Beweis:

Der Beweis gelingt durch vollständige Induktion über n:

Induktionsanfang ($n = 0$):
Es gelte $r_\sigma \in \langle t \rangle_A^0 = \{t_{id}\}$. Dann folgt

$$r\tau = t\tau \in [t\tau]_{(A,\tau)}^*$$

und daraus nach Definition 4.6.1

$$(r\tau)_\tau \in \langle t \rangle_A.$$

Induktionsschluß ($n \to n+1$):
Sei $r_\sigma \in \langle t \rangle_A^{n+1}$. Dann sind zwei Fälle zu unterscheiden, nämlich $r_\sigma \in \langle t \rangle_A^n$ und $r_\sigma \in \Theta_n$.

Wenn $r_\sigma \in \langle t \rangle_A^n$, dann ist die Induktionsannahme direkt anwendbar, und es folgt $(r\tau)_\tau \in \langle t \rangle_A$.

Ist $r_\sigma \in \Theta_n$, so gibt es ein Element $t'_{\tau'} \in \langle t \rangle_A^n$ aus dem r_σ in einem Schritt ableitbar ist. Mit Hilfe einer Gleichung $G \in \mathrm{Gl}^*(A, \sigma)$ ist also der Term r aus $t'\sigma$ ableitbar, und es gilt $\tau' \succ \sigma$ und daher auch $\tau' \succ \tau$.

Auf $t'_{\tau'} \in \langle t \rangle_A^n$ und τ ist nun die Induktionsannahme anwendbar, und es folgt $(t'\tau)_\tau \in \langle t \rangle_A$ und daraus

$$t'\tau \in [t\tau]_{(A,\tau)}^*.$$

Da τ eine Spezialisierung von σ ist, ist $r\tau$ aus $t'\tau$ mit Hilfe der Gleichung $G\tau \in \mathrm{Gl}^*(A,\tau)$ ableitbar. Darum liegt auch $r\tau$ in $[t\tau]^*_{(A,\tau)}$, woraus, wiederum nach Definition 4.6.1,

$$(r\tau)_\tau \in \langle t\rangle_A$$

folgt. ■

Lemma 4.7.12

Es gelte

$$r_\sigma \in \langle t\rangle_A.$$

Dann gibt es für ein $n \geq 0$ ein Element

$$r'_{\sigma'} \in \langle t\rangle_A^n$$

mit $r'_{\sigma'} \gg r_\sigma$.

Beweis:

Es gelte

$$r_\sigma \in \langle t\rangle_A,$$

und daher

$$r \in [t\sigma]^*_{(A,\sigma)}.$$

Dann gibt es für ein $m \geq 0$ eine Folge $G_1, \ldots, G_m$ von Gleichungen aus $\mathrm{Gl}^*(A,\sigma)$, mit deren Hilfe r aus $t\sigma$ über die Zwischenstufen $u^0, \ldots, u^m$ in folgender Weise ableitbar ist:

$$t\sigma = u^0 \stackrel{G_1}{\rightarrow} u^1 \stackrel{G_2}{\rightarrow} \cdots \stackrel{G_m}{\rightarrow} u^m = r$$

Durch vollständige Induktion wird für $0 \leq k \leq m$ gezeigt:

Induktionsannahme ($k = 0$):

Das einzige Element $r^0_{\sigma_0} = t_{id} \in \langle t\rangle_A^0$ subsumiert $u^0_\sigma = (t\sigma)_\sigma$.

Induktionsanfang:

Für ein $l_k \geq 0$ gibt es ein Element

$$r^k_{\sigma_k} \in \langle t\rangle_A^{l_k}$$

mit $r^k_{\sigma_k} \gg u^k_\sigma$.

Induktionsschluß ($k \to k+1$):

Aus der Induktionsannahme folgt, daß es für ein $l_k \geq 0$ ein Element

$$r^k_{\sigma_k} \in \langle t\rangle_A^{l_k}$$

gibt mit $r^k_{\sigma_k} \gg u^k_\sigma$.

Da die Auswahlheuristik $\mathcal{H}$ fair ist, muß schließlich ein $r'^k_{\sigma'_k}$ ausgewählt werden, daß $r^k_{\sigma_k}$ subsumiert. Es muß also ein $l_{k+1} \geq 0$ geben mit

$$\mathcal{H}(\langle t\rangle_A^{l_{k+1}-1}) = r'^k_{\sigma'_k} \gg r^k_{\sigma_k} \gg u^k_\sigma.$$

Da σ eine Grundsubstitution ist und u^k ein geschlossener Term, folgt daraus

$$r'^k\sigma = r^k\sigma = u^k\sigma = u^k \stackrel{G_{k+1}}{\rightarrow} u^{k+1}.$$

Außerdem gilt $\sigma'_k \succ \sigma$. Damit muß es nach Definition 4.7.4 eine Substiution σ'_{k+1} geben, die σ subsumiert, und mit der

$$u^{k+1}_{\sigma'_{k+1}} \in \Theta_{k+1}$$

gilt. Dann gibt es aber auch ein Element

$$r^{k+1}_{\sigma'_{k+1}} \in \langle t\rangle_A^{k+1},$$

das $u^{k+1}_{\sigma'_{k+1}}$ und daher auch u^{k+1}_σ subsumiert.

Gestützt auf diesen Induktionsbeweis erhält man nun für $k = m$: Es gibt ein

$$r'_{\sigma'} = r^m_{\sigma_m} \in \langle t\rangle_A^n = \langle t\rangle_A^{l_m}$$

mit $r'_{\sigma'} \gg u^m_\sigma = r_\sigma$. ■

Mit Hilfe der beiden gerade bewiesenen Lemmata kann nun der Abschluß nach Definition 4.7.10 auf denjenigen nach Definition 4.6.2 zurückgeführt werden, dessen Korrektheit und Vollständigkeit schon bewiesen ist (Satz 4.6.4).

Satz 4.7.13

Sei T das nach Definition 4.3.6 unter Beachtung der Grenze q für die Anwendung der γ–Regel aufgebaute Tableau für eine Formel X.

Korrektheit: Ist T nach Definition 4.7.10 geschlossen, dann ist die Formel X unerfüllbar.

Vollständigkeit: Ist X unerfüllbar und die Grenze q hinreichend hoch gewählt, dann ist T nach Definition 4.7.10 geschlossen.

Beweis:

Korrektheit:
T sei nach Definition 4.7.10 geschlossen.

Die Grundsubstitution τ sei eine Spezialisierung der nach Definition 4.7.10 existierenden Substitution σ und damit auch jeder der Substitutionen in

$$\{\rho_{ij}, \overline{\rho}_{ij}, \overline{\overline{\rho}}_{ij} \ : \ 1 \leq i \leq k,\, 1 \leq j \leq n_i\},$$

deren Vereinigung σ ist. Gibt es mehrere solcher Grundsubstitutionen, sei eine beliebig aber fest ausgewählt.

Für beliebige $1 \leq i \leq k$ und $1 \leq j \leq n_i$ gilt

$$r^{ij}_{\rho_{ij}} \in \langle t^*_{ij} \rangle^{l_{ij}}_{A_i} \quad \text{und} \quad \overline{r}^{ij}_{\overline{\rho}_{ij}} \in \langle s^*_{ij} \rangle^{l_{ij}}_{A_i}.$$

Daraus folgt mit Lemma 4.7.11

$$(r^{ij}\tau)_\tau \in \langle t^*_{ij} \rangle_{A_i} \quad \text{und} \quad (\overline{r}^{ij}\tau)_\tau \in \langle s^*_{ij} \rangle_{A_i}.$$

Damit haben aber, da τ eine Spezialisierung des Unifikators von r^{ij} und $\overline{r}^{ij}$ ist, und also $r^{ij}\tau = \overline{r}^{ij}\tau$ gilt, die Mengen $\langle t^*_{ij} \rangle_{A_i}$ und $\langle s^*_{ij} \rangle_{A_i}$ ein gemeinsames Element.

Also ist T auch nach Definition 4.6.2 geschlossen, und daher ist X nach Satz 4.6.4 unerfüllbar.

Vollständigkeit:
Sei X unerfüllbar und q hinreichend hoch gewählt. Dann ist T nach Definition 4.6.2 geschlossen (Satz 4.6.4).

Daraus folgt zunächst für beliebige $1 \leq i \leq k$ und $1 \leq j \leq n_i$ die Existenz von Elementen

$$\hat{r}^{ij}_\sigma \in \langle t^*_{ij} \rangle_{A_i} \cap \langle s^*_{ij} \rangle_{A_i}$$

(diese sind in Definition 4.6.2 mit r^{ij}_σ bezeichnet).

Nach Lemma 4.7.12 gibt es dann Elemente

$$r_{\rho_{ij}} \in \langle t^*_{ij} \rangle^{l_{ij}}_{A_i} \quad \text{und} \quad \overline{r}_{\overline{\rho}_{ij}} \in \langle t^*_{ij} \rangle^{\overline{l}_{ij}}_{A_i}$$

mit

$$r_{\rho_{ij}} \gg \hat{r}^{ij}_\sigma \quad \text{und} \quad \overline{r}_{\overline{\rho}_{ij}} \gg \hat{r}^{ij}_\sigma.$$

Es gibt also Substitutionen μ und ν, so daß

$$r^{ij}\mu = \overline{r}^{ij}\nu = \hat{r},$$

und die Substitutionen μ, ν, ρ_{ij} und $\overline{\rho}_{ij}$ allgemeiner sind als die Substitution σ.

Da μ und ν allgemeiner als σ sind, gilt $r^{ij}\sigma = \overline{r}^{ij}\sigma$. r^{ij} und $\overline{r}^{ij}$ haben also einen MGU $\overline{\overline{\rho}}_{ij}$, der allgemeiner als σ ist.

Da nun ρ_{ij}, $\overline{\rho}_{ij}$ und $\overline{\overline{\rho}}_{ij}$ die geforderten Eigenschaften besitzen, ist T auch nach Definition 4.7.10 geschlossen. ∎

4.8 Steuerung der Suche nach dem Abschluß eines Tableaus

Die Überprüfung darauf, ob ein Tableau nach Definition 4.7.10 geschlossen ist, kann, da die iterative Definition von $\langle A\rangle_t^n$ zur Verfügung steht, nun in folgender Weise erfolgen und implementiert werden.

Definition 4.8.1 (Verfahren zum Auffinden eines Abschlusses)

Ein Tableau T mit den Ästen $A_1, \ldots, A_k$ sei vorgelegt, das nach Definition 4.3.6 aufgebaut und also unter Beachtung einer Grenze q für die Anwendung der γ–Regel voll expandiert ist.

Nach folgendem Algorithmus kann überprüft werden, ob es nach Definition 4.7.10 geschlossen ist:

1. *Zunächst werden die Mengen* $\mathrm{Gl}^*(A_i, id)$ *und*

   $$\mathrm{Dis}^*(A_i) = \{D_1^{*i}, \ldots, D_{m_i}^{*i}\}$$

   für $1 \leq i \leq k$ bestimmt. Dabei ist

   $$D_j^{*i} = (t_{j1}^{*i} \not\approx s_{j1}^{*i} \vee \ldots \vee t_{jp_{ij}}^{*i} \not\approx s_{jp_{ij}}^{*i}).$$

2. *Für jeden der Äste A_i legt man eine Menge Γ^i von ihn schließenden Substitutionen an. Das gleiche geschieht für jeden der in einer der Mengen* $\mathrm{Dis}^*(A_i)$ $(1 \leq i \leq k)$ *enthaltenen Terme. In Γ_{jl}^i sind also Substitutionen enthalten, die es erlauben, die Ungleichung $t_{jl}^{*i} \not\approx s_{jl}^{*i}$ zum Widerspruch zu führen. Alle diese Mengen werden zu Anfang als leer initialisiert.*

3. *Nun führt man einen Iterationsschritt aus, um für t_{jl}^{*i} und s_{jl}^{*i} die Mengen $\langle t_{jl}^{*i}\rangle_{A_i}^{n+1}$ bzw. $\langle s_{jl}^{*i}\rangle_{A_i}^{n+1}$ zu berechnen und zwar für alle $1 \leq i \leq k$, $1 \leq j \leq m_i$ und $1 \leq l \leq p_{ij}$, also für alle in einer der Menge* $\mathrm{Dis}^*(A_i)$ $(1 \leq i \leq k)$ *auftretenden Terme.*

   Treten dabei keine neuen Elemente mehr auf, und ist noch kein Abschluß des Tableaus gefunden, scheitert das Verfahren und das Tableau kann nicht geschlossen werden.

4. *Treten für ein n Elemente*

   $$r_\rho \in \langle t_{jl}^{*i}\rangle_{A_i}^n \quad \text{und} \quad \overline{r}_{\overline{\rho}} \in \langle s_{jl}^{*i}\rangle_{A_i}^n$$

   auf, so daß r und $\overline{r}$ unifizierbar sind mit einem MGU $\overline{\overline{\rho}}$, und ρ, $\overline{\rho}$ und $\overline{\overline{\rho}}$ eine kanonische Vereinigung σ besitzen, dann kann mit dieser Substitution σ die Ungleichung widerlegt werden, und man fügt σ zur Menge Γ_{jl}^i hinzu — es sei denn, es gibt schon eine Substitution in Γ_{jl}^i, die allgemeiner ist als σ.

5. *Hat man gemäß 4. zu einer Menge Γ_{jl}^i neue Substitutionen hinzugefügt, so überprüft man, ob es nun Substitutionen*

   $$\sigma_1^i \in \Gamma_{j1}^i, \ \ldots, \ \sigma_{p_{ij}}^i \in \Gamma_{jp_{ij}}^i$$

   gibt, die eine kanonische Vereinigung σ besitzen. Ist dies der Fall, kann die Disjunktion D_j^{*i} mit σ widerlegt und damit der Ast A_i geschlossen werden. Man fügt also σ zu Γ^i hinzu — unter der Bedingung, daß es nicht schon eine Substitution in Γ^i gibt, die allgemeiner ist als σ.

6. *Ist eine Menge Γ^i nach 5. erweitert worden, so überprüft man, ob es nun Substitutionen*

   $$\sigma_1 \in \Gamma^1 \ \ldots \ \sigma_k \in \Gamma^k$$

   gibt, die eine kanonische Vereinigung σ besitzen.

   Hat man solche Substitutionen gefunden, ist damit das Tableau geschlossen, und der Algorithmus terminiert.

   Andernfalls wird gemäß 3. mit der iterativen Berechnung der $\langle t_{jl}^{*i}\rangle_{A_i}^{n+1}$ und $\langle s_{jl}^{*i}\rangle_{A_i}^{n+1}$ fortgefahren.

Es ist nicht erforderlich, daß die Berechnung der $\langle t_{jl}^{*i}\rangle_{A_i}^n$ und $\langle s_{jl}^{*i}\rangle_{A_i}^n$ gleichmäßig vorangetrieben wird. Es muß nur sichergestellt sein, daß jede der Mengen schließlich für jedes $n \geq 0$ bestimmt wird.

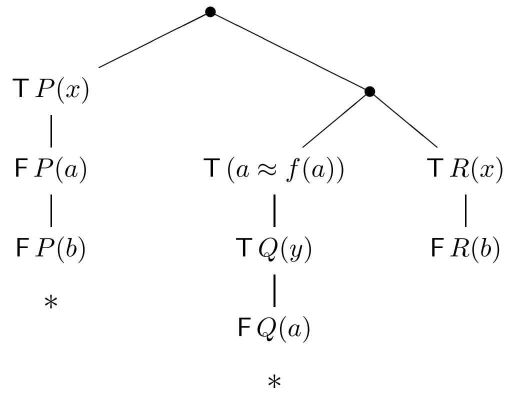


Abbildung 4.7: Ausschnitt aus einem Tableau, dessen Abschluß ohne eine Einschränkung der Vollständigkeit nicht gezeigt werden kann.

Um das Verfahren genau so, wie es in Definition 4.8.1 beschrieben ist, zu implementieren, hätte die globale Ablaufsteuerung von $_3T^AP$ verändert werden müssen, was aber nach Möglichkeit vermieden werden sollte. Bisher wurden von $_3T^AP$ die Äste eines Tableaus der Reihe nach geschlossen. Die zum Abschluß eines Astes notwendigen Variablenbelegungen wurden ausgeführt, bevor die weiteren Äste betrachtet wurden. Die Suche nach einer Substitution, die den Abschluß aller Äste des Tableaus gestattet, wurde dem Backtracking des PROLOG überlassen.

Um diese Ablaufsteuerung im wesentlichen beibehalten zu können, wird das in Definition 4.8.1 beschriebene Verfahren dahingehend abgeändert, daß nur noch jeweils ein Ast betrachtet, und eine ihn abschließende Substitution gesucht wird.

Die einen Ast abschließende Substitution wird sofort ausgeführt, und anschließend erst werden die weiteren Äste des Tableaus betrachtet. Ist es dann nicht möglich, das ganze Tableau zu schließen, wird es dem Backtracking überlassen, für schon geschlossene Äste andere schließende Substitutionen zu finden.

Mit dieser Änderung geht zunächst die Vollständigkeit des Kalküls verloren, wie folgendes Beispiel zeigt:

Beispiel 4.8.2

In Abbildung 4.7 ist ein Tableau dargestellt, dessen Abschluß mit dem abgeänderten Verfahren, das auf der Betrachtung jeweils nur eines Astes beruht, nicht gezeigt werden kann.

Ist nämlich der linke Ast des Tableaus mit den Atomen T $P(x)$ und F $P(a)$ und der Substitution $\{x \leftarrow a\}$ geschlossen, dann läßt sich der rechte Ast nicht mehr schließen, wenn nicht die beiden Vorkommen des Atoms T $P(x)$ als universell erkannt werden. Das ist aber nur schwer möglich, jedenfalls nicht nach Satz 3.1.3.

Wird nun Backtracking ausgelöst, weil der rechte Ast nicht geschlossen werden kann, dann wird zunächst ein anderer Abschluß des mittleren Astes gesucht. Der mittlere Ast hat aber unendlich viele verschiedene Abschlüsse $\{y \leftarrow a\}$, $\{y \leftarrow f(a)\}$ usw., die der Reihe nach bestimmt werden. Zu erkennen, daß diese anderen Abschlüsse des mittleren Astes das Problem im rechten Ast nicht lösen, sondern daß der Abschluß des linken Astes geändert werden muß, ist in diesem speziellen Fall vielleicht noch möglich. Im allgemeinen ist es aber ein unentscheidbares Problem.

Da $_3T^AP$s globale Ablaufsteuerung beibehalten und daher von der Betrachtung nur jeweils eines Astes nicht abgewichen werden soll, bietet sich nur folgende Lösung an: Die Vollständigkeit des Verfahrens wird — wie schon durch die Einführung der Grenze q für die Anwendung der γ–Regel — in sinnvoller Weise eingeschränkt. Die maximale Länge der Ableitungsketten, die zum Abschluß eines Tableaus betrachtet werden, wird begrenzt durch eine Schranke r.

Das Verfahren zum Auffinden eines Abschlusses wird also in folgender Weise abgeändert:

Definition 4.8.3 (Implementiertes Verfahren)

Wie in Definition 4.8.1 sei ein Tableau T mit den Ästen $A_1, \dots, A_k$ vorgelegt, das unter Beachtung einer Grenze q für die Anwendung der γ–Regel voll expandiert ist.

Um zu prüfen, ob das Tableau nach Definition 4.7.10 geschlossen ist, werden nun aber die Äste A_i beginnend mit $i = 1$ einzeln der Reihe nach betrachtet:

1. *Zunächst werden die Mengen $\mathrm{Gl}^*(A_i, id)$ und*

$$\mathrm{Dis}^*(A_i) = \{D_1^*, \dots, D_m^*\}$$

   bestimmt. Dabei ist

$$D_j^* = (t_{j1}^* \not\approx s_{j1}^* \vee \dots \vee t_{jp_j}^* \not\approx s_{jp_j}^*).$$

2. *Für jede Ungleichung $t_{jl}^* \not\approx s_{jl}^*$ legt man eine Menge Γ_{jl} der sie schließenden Substitutionen an und initialisiert sie als leer.*

3. *Anschließend führt man einen Iterationsschritt aus, um für t_{jl}^* und s_{jl}^* die Mengen $\langle t_{jl}^* \rangle_{A_i}^{n+1}$ bzw. $\langle s_{jl}^* \rangle_{A_i}^{n+1}$ zu berechnen und zwar für alle $1 \leq j \leq m_i$ und $1 \leq l \leq p_{ij}$, also für alle in $\mathrm{Dis}(A_i)$ auftretenden Terme.*
   Treten dabei keine neuen Elemente mehr auf, ohne daß die Grenze r für die Länge von Ableitungsketten überschritten wird, und ist noch kein Abschluß des Tableaus gefunden, muß Backtracking eingeleitet, der vorangehende Ast A_{i-1} erneut betrachtet und die Iteration für diesen Ast fortgeführt werden, um weitere schließende Substitutionen zu finden.
   Wenn $i = 1$, scheitert das Verfahren und das Tableau kann nicht geschlossen werden.

4. *Treten für ein n Elemente*

$$r_\rho \in \langle t_{jl}^* \rangle_{A_i}^n \text{ und } \overline{r}_{\overline{\rho}} \in \langle s_{jl}^* \rangle_{A_i}^n$$

   auf, so daß r und $\overline{r}$ unifizierbar sind mit einem MGU $\overline{\overline{\rho}}$, und ρ, $\overline{\rho}$ und $\overline{\overline{\rho}}$ eine kanonische Vereinigung σ besitzen, dann kann mit dieser Substitution σ die Ungleichung widerlegt werden.
   Man fügt σ zur Menge Γ_{jl} hinzu — es sei denn, es gibt schon eine Substitution in Γ_{jl}, die allgemeiner ist als σ.

5. *Hat man gemäß 4. zu einer Menge Γ_{jl} neue Substitutionen hinzugefügt, so überprüft man, ob es nun Substitutionen*

$$\sigma_1 \in \Gamma_{j1}, \dots, \sigma_{p_j} \in \Gamma_{jp_j}$$

   gibt, die eine kanonische Vereinigung σ besitzen. Ist dies der Fall, kann die Disjunktion D_j^* mit dieser Substitution σ widerlegt und damit der Ast A_i geschlossen werden. Man führt also die Substitution σ aus und fährt mit dem nächsten Ast fort. Gibt es keinen weiteren Ast, ist also $i = k$, so ist das Tableau geschlossen und das Verfahren beendet.
   Andernfalls wird gemäß 3. mit der iterativen Berechnung der $\langle t_{jl}^* \rangle_{A_i}^n$ und $\langle s_{jl}^* \rangle_{A_i}^n$ fortgefahren.

Es ist nicht erforderlich, daß die Berechnung der $\langle t_{jl}^* \rangle_{A_i}^n$ und $\langle s_{jl}^* \rangle_{A_i}^n$ gleichmäßig vorangetrieben wird. Es muß nur sichergestellt sein, daß jede der Mengen schließlich für jedes $n \geq 0$ bestimmt wird.

Es tritt nun das schon in Abschnitt 4.5 erläuterte Problem der Tiefensuche auf. Denn eine Tiefensuche wird nun wieder — durch Backtracking modelliert — für die Suche nach einer Substitution ausgeführt, die nicht nur den Abschluß einzelner sondern aller Äste gestattet. Man kann zwar durch eine geschickte Wahl der Grenze r regelnd eingreifen. Andererseits stellt die Wahl von r aber auch selbst wieder ein Problem dar.

Hat man r zu groß gewählt, gelingt der Abschluß in Fällen, wie dem in Beispiel 4.8.2 dargestellten nicht — zumindest nicht in angemessener Zeit. Wählt man andererseits r zu klein, gelingen zu viele Beweise überhaupt nicht mehr, da für den Beweis eine längere Ableitungskette nötig wäre. In der Praxis haben sich aber Werte um $r = 10$ für die meisten Probleme als geeignet erwiesen.

$$\mathsf{F}\,(g(f(a)) \approx g(a))$$
$$\mathsf{T}\,(\forall x)(\neg(x \approx a) \vee g(f(x)) \approx g(x))$$
$$\mathsf{T}\,(\neg(x_1 \approx a) \vee g(f(x_1)) \approx g(x_1))$$
$$\begin{array}{ll} \mathsf{T}\,\neg(x_1 \approx a) & \qquad \mathsf{T}\,(g(f(x_1)) \approx g(x_1)) \\ \mathsf{F}\,(x_1 \approx a) & \end{array}$$

Abbildung 4.8: Ein Tableau, dessen Abschluß mit Hilfe des Verfahrens aus Definition 4.8.3 nicht gezeigt werden kann (Beispiel 4.8.4).

Zu zeigen bleibt, daß Korrektheit und Vollständigkeit des Kalküls nicht dadurch verloren gehen, daß, direkt nachdem der Abschluß eines Astes gefunden wurde, die den Abschluß zulassende Substitution auf das ganze Tableau angewendet wird, also noch bevor die Abschlüsse der weiteren Äste bestimmt sind.

Wie man erwartet, gilt dies für die Korrektheit ohne jede Einschränkung: Ist T ein Tableau für eine Formel X, und wird der Abschluß von T festgestellt, so ist X unerfüllbar, und T nach allen zuvor gegebenen Definitionen für den Abschluß eines Tableaus geschlossen.

Mit der Vollständigkeit verhält es sich jedoch etwas anders. Sie kann, wie Beispiel 4.8.4 zeigt, nicht in der Weise formuliert werden, daß der Abschluß jedes nach Definition 4.7.10 geschlossenen Tableaus auch durch das veränderte Verfahren nach Definition 4.8.3 festgestellt wird. Der Grund dafür ist, daß unter Umständen die Universalität einer Formel nicht mehr ausgenutzt werden kann, nachdem eine Substitution auf das Tableau angewendet wurde. In solchen Fällen kann der Abschluß des Tableaus durch das Verfahren aus Definition 4.8.3 nur dann gezeigt werden, wenn die Grenze q für die Anwendung der γ–Regel beim Aufbau von T hinreichend hoch — höher als es für den Abschluß nach Definition 4.7.10 notwendig wäre — gewählt wurde, denn dann fällt der Verlust der Universalität von Formeln nicht mehr ins Gewicht.

Setzt man allerdings die — für die Implementierung verwendete — Methode aus Satz 3.1.3 zum Auffinden universeller Formeln ein, so kommt die gerade geschilderte Problematik nicht zum tragen. Denn dann treten Variablen, bezüglich derer Formeln als universell erkannt werden, immer nur in einem Ast auf und werden daher nicht durch den Abschluß anderer Äste belegt.

Beispiel 4.8.4

Das in Abbildung 4.8 dargestellte Tableau ist nach Definition 4.3.6 unter Beachtung der Grenze $q = 1$ für die Anwendung der γ–Regel aufgebaut und nach Definition 4.7.10 geschlossen — vorausgesetzt, man hat eine Methode zur Verfügung, mit deren Hilfe es gelingt, die Formel $\mathsf{T}\,(g(f(x_1)) \approx g(x_1))$ als universell bezüglich x_1 zu erkennen (die Methode aus Satz 3.1.3 leistet dies nicht).

Setzt man nun das Verfahren aus Definition 4.8.3 ein, um zu zeigen, daß das Tableau geschlossen ist, so wird zunächst der linke Ast mit Hilfe der Substitution $\{x \leftarrow a\}$ geschlossen, und diese auf das ganze Tableau angewendet. Dann gelingt es aber nicht mehr, den rechten Ast des Tableaus zu schließen, denn die nun entstandene Gleichung $\mathsf{T}\,(g(f(a)) \approx g(a))$ genügt dafür nicht.

Satz 4.8.5

Sei T das nach Definition 4.3.6 unter Beachtung der Grenze q für die Anwendung der γ–Regel aufgebaute Tableau für eine Formel X.

Korrektheit: Wird der Abschluß von T durch das Verfahren aus Definition 4.8.3 festgestellt, dann ist die Formel X unerfüllbar.

Vollständigkeit: Ist die Formel X unerfüllbar, und sind q und die Grenze r für die Länge von Ableitungsketten hinreichend hoch gewählt, dann wird der Abschluß von T durch das in Definition 4.8.3 beschriebene Verfahren festgestellt.

Beweis:

Korrektheit:
Der Abschluß des Tableaus T sei durch das Verfahren aus Definition 4.8.3 festgestellt worden.

Für $1 \leq i \leq k$ sei mit τ_i diejenige kanonische Substitution bezeichnet, die zum Abschluß des Astes A_i auf das Tableau angewendet und nicht wieder durch Backtracking verworfen wurde.

Die Substitutionen σ_i seien definiert durch

$$\sigma_0 = id \text{ und } \sigma_i = \tau_i \circ \cdots \circ \tau_1.$$

Dann ist $T\sigma_i$ das Tableau, das nach Abschluß des i–ten Astes und Anwendung der Substitution τ_i bzw. vor Abschluß des $(i+1)$–ten Astes vorliegt.

Die Substitutionen τ_i' seien dadurch gebildet, daß man die Belegungen derjenigen Variablen, die in $T\sigma_{i-1}$ und in $(\mathrm{Dis}^*(A_m))\sigma_{i-1}$ $(1 \leq m \leq k)$ nicht frei auftreten, aus τ_i entfernt:

$$x\tau_i' = \begin{cases} x\tau_i & x \text{ tritt in } T\sigma_{i-1} \text{ oder für ein } 1 \leq m \leq k \text{ in } (\mathrm{Dis}^*(A_m))\sigma_{i-1} \text{ frei auf} \\ x & \text{sonst} \end{cases}$$

Analog zu σ_i sei σ_i' gegeben durch

$$\sigma_0' = id \text{ und } \sigma_i' = \tau_i' \circ \cdots \circ \tau_1'.$$

Damit folgt für alle $1 \leq i \leq k$

$$T\sigma_i = T\sigma_i'.$$

Aufgrund der Bedingungen in den Schritten 4 und 5 des Verfahrens gibt es für $1 \leq i \leq k$ Disjunktionen

$$D_i^* = (t_{i1}^* \not\approx s_{i1}^* \vee \cdots \vee t_{in_i}^* \not\approx s_{in_i}^*),$$

so daß für $1 \leq j \leq n_i$ Elemente

$$r_{\rho_{ij}}^{ij} \in \langle t_{ij}^* \sigma_{i-1}' \rangle_{A_i\sigma_{i-1}'}^{l_{ij}} \text{ und } \overline{r}_{\overline{\rho}_{ij}}^{ij} \in \langle s_{ij}^* \sigma_{i-1}' \rangle_{A_i\sigma_{i-1}'}^{\overline{l}_{ij}}$$

existieren, und r^{ij} und $\overline{r}^{ij}$ unifizierbar sind mit einem MGU $\overline{\overline{\rho}}_{ij}$, und die Substitution τ_i eine Spezialisierung ist von jeder der Substitutionen in

$$\{\rho_{ij}, \overline{\rho}_{ij}, \overline{\overline{\rho}}_{ij} \;:\; 1 \leq i \leq k, 1 \leq j \leq n_i\}.$$

ν sei eine beliebige Grundsubstitution, die eine Spezialisierung von σ_k' und damit auch von $\sigma_1' \succ \cdots \succ \sigma_{k-1}' \succ \sigma_k'$ ist.

Die Substitutionen ν_i seien durch

$$\nu_i = \nu \circ \tau_i$$

definiert. Sie sind Grundsubstitutionen, und es gilt $\tau_i \succ \nu_i$.

τ_i' belegt keine Variablen, die durch σ_{i-1}' belegt werden, denn solche treten in $T\sigma_{i-1} = T\sigma_{i-1}'$ und in $(\mathrm{Dis}^*(A_i))\sigma_{i-1} = (\mathrm{Dis}^*(A_i))\sigma_{i-1}'$ nicht mehr frei auf. Darum gibt es eine Belegung μ_i, so daß $\sigma_i' = \tau_i' \circ \sigma_{i-1}' = \mu_i \circ \tau_i'$, woraus $\tau_i' \succ \sigma_i'$ und also $\tau_i' \succ \tau$ folgt.

Ohne Beschränkung der Allgemeinheit kann man davon ausgehen, daß Variablen, die in $T\sigma_i$ frei auftreten, in $T\sigma_{i-1}$ aber nicht, die also durch die Anwendung von τ_i neu eingeführt wurden, auch in $T\sigma_0, \ldots, T\sigma_{i-2}$ nicht frei auftreten, also tatsächlich völlig neu sind. Notfalls kann man dies durch Umbenennung der Variablen in T und in den Abschlüssen der Äste $A_1, \ldots, A_{i-2}$ erreichen. Dies hat zur Folge, daß nicht nur die Substitutionen τ_i und τ_i', sondern auch σ_i und σ_i' für $1 \leq i \leq k$ kanonisch sind.

Für $1 \leq i \leq k$ und $1 \leq j \leq n_i$ gilt nun $\rho_{ij} \succ \tau_i \succ \nu_i$ und analog $\overline{\rho}_{ij} \succ \nu_i$ und $\overline{\overline{\rho}}_{ij} \succ \nu_i$. Damit folgt nun aus Lemma 4.7.11

$$(r^{ij}\nu_i)_{\nu_i} \in \langle t_{ij}^* \sigma_{i-1} \rangle_{A_i\sigma_{i-1}} \text{ und } (\overline{r}\nu_i^{ij})_{\nu_i} \in \langle s_{ij}^* \sigma_{i-1} \rangle_{A_i\sigma_{i-1}}$$

und daraus

$$[t_{ij}^* \sigma_{i-1} \nu_i]^*_{(A_i\sigma_{i-1}, \nu_i)} = [s_{ij}^* \sigma_{i-1} \nu_i]^*_{(A_i\sigma_{i-1}, \nu_i)}.$$

Nun ist aber

$$t^*_{ij}\sigma_{i-1}\nu_i = t^*_{ij}\sigma_{i-1}\tau_i\nu,$$

und da sich τ_i und τ'_i nach Definition nur in Variablen unterscheiden können, die in $t^*_{ij}\sigma_{i-1}$ nicht frei auftreten, gilt weiter

$$t^*_{ij}\sigma_{i-1}\nu_i = t^*_{ij}\sigma_{i-1}\tau_i\nu = t^*_{ij}\sigma_{i-1}\tau'_i\nu.$$

Die Substitutionen σ_{i-1} und τ'_i sind kanonisch und allgemeiner als ν. Darum ist $\sigma_{i-1}\tau'_i\nu = \nu$. Schließlich erhält man also

$$t^*_{ij}\sigma_{i-1}\nu_i = t^*_{ij}\sigma_{i-1}\tau_i\nu = t^*_{ij}\sigma_{i-1}\tau'_i\nu = t^*_{ij}\nu.$$

und analog

$$s^*_{ij}\sigma_{i-1}\nu_i = s^*_{ij}\nu.$$

Wiederum gestützt auf das Argument, daß sich τ_i und τ'_i nur in Variablen unterscheiden können, die in $T\sigma_{i-1}$ nicht frei auftreten, erhält man $\mathrm{Gl}^*(A_i\sigma_{i-1}, \nu_i) = \mathrm{Gl}^*(A_i\sigma'_{i-1}, \nu)$. Zusammengefaßt ergibt sich

$$[t^*_{ij}\nu]^*_{(A_i\sigma'_{i-1},\nu)} = [s^*_{ij}\nu]^*_{(A_i\sigma'_{i-1},\nu)}.$$

Da ν eine Spezialisierung von σ'_{i-1} ist, gilt dann erst recht

$$[t^*_{ij}\nu]^*_{(A_i,\nu)} = [s^*_{ij}\nu]^*_{(A_i,\nu)},$$

und das Tableau T ist nach Definition 4.4.6 geschlossen, woraus mit Lemma 4.4.9 die Unerfüllbarkeit der Formel X folgt.

Vollständigkeit:
X sei unerfüllbar. Mit $\mathcal{M}_0$ sei die triviale Methode zum Auffinden universeller Formeln bezeichnet, die keine universellen Formeln auffindet.

Die Grenze q für die Anwendung der γ–Regel sei so hoch gewählt, daß das unter Beachtung dieser Grenze aufgebaute Tableau T für X nach Definition 4.7.10 unter Verwendung von $\mathcal{M}_0$ geschlossen ist, was nach Satz 4.7.13 möglich ist.

Es wird nun ein Induktionsbeweis geführt und gezeigt, daß, wenn ein solches Tableau T mit den Ästen $A_1, \ldots, A_k$ dem Verfahren vorgelegt wird, folgendes für alle $0 \le n \le k$ gilt:

Induktionsannahme
Nach endlich vielen Ableitungsschritten und ggf. nachdem andere Abschlüsse durch Backtracking verworfen wurden, tritt der Zustand ein, daß schon n Äste des Tableaus in Schritt 5 des Verfahrens geschlossen wurden, und daß das dann vorliegende Tableau T_n wieder nach Definition 4.7.10 geschlossen ist.

Da diese Aussage dann insbesondere auch für $n = k$ gilt, wird der Abschluß aller k Äste des Tableaus schließlich durch das Verfahren festgestellt.

Induktionsanfang ($n = 0$):
Sind $n = 0$ Äste geschlossen, liegt das ursprüngliche Tableau $T = T_0$ vor. Es ist als nach Definition 4.7.10 geschlossen vorausgesetzt.

Induktionsschluß ($n \to n + 1$):
Die Äste des Tableaus T_n seien mit $A^n_1, \ldots, A^n_k$ bezeichnet, wobei für $1 \le i \le k$ der Ast A^n_i aus dem Ast A_i des ursprünglichen Tableaus T hervorgegangen ist.

Da T_n nach Definition 4.7.10 geschlossen ist (und zwar unter Verwendung von $\mathcal{M}_0$), gibt es für $1 \le i \le k$ Disjunktionen

$$D^*_i = (t^*_{i1} \not\approx s^*_{i1} \vee \cdots \vee t^*_{ip_i} \not\approx s^*_{ip_i}) \in \mathrm{Dis}^*(A^n_i),$$

so daß es für $1 \le j \le p_i$ Elemente

$$r^{ij}_\rho \in \langle t^*_{ij}\rangle^{l_{ij}}_{A^n_i} \quad \text{und} \quad \overline{r}^{ij}_{\overline{\rho}} \in \langle s^*_{ij}\rangle^{\overline{l}_{ij}}_{A^n_i},$$

gibt, und r^{ij} und $\overline{r}^{ij}$ unifizierbar sind mit einem MGU $\overline{\overline{\rho}}_{ij}$. Außerdem besitzen die Substitutionen in

$$\{\rho_{(n+1)j}, \overline{\rho}_{(n+1)j}, \overline{\overline{\rho}}_{(n+1)j} \;:\; 1 \le j \le p_i\}$$

eine kanonische Vereinigung τ. Diese kann allgemeiner sein als die Substitution σ, deren Existenz in Definition 4.7.10 gefordert wird, denn σ ist die kanonische Vereinigung einer umfassenderen Menge von Substitutionen.

Es ist damit sichergestellt, daß die Suche nach einem Abschluß des Astes A^n_{n+1} erfolgreich ist (die Grenze r sei hinreichend hoch gewählt). Es kann aber sein, daß zunächst noch andere Abschlüsse gefunden werden. Wurde ein anderer Abschluß gewählt, und gelingt es danach, den Abschluß aller weiteren Äste festzustellen, so folgt aus der schon bewiesenen Korrektheit des Verfahrens, daß die entstehenden Tableaus $T_{n+1}, \ldots, T_k$ nach Definition 4.7.10 geschlossen sind. Viel wichtiger ist aber, daß dann der Abschluß aller Äste gelungen ist.

Andernfalls, wenn die weiteren Äste nicht geschlossen werden können, tritt Backtracking auf. Da die Suche nach Abschlüssen durch r begrenzt ist (in endlich vielen Schritten können nur endlich viele Terme abgeleitet werden), muß schließlich der Abschluß mit der auszuführenden Substitution τ aufgefunden werden, dessen Existenz sichergestellt ist. Es bleibt zu zeigen, daß das Tableau $T_{n+1} = T_n\tau$ wieder nach Definition 4.7.10 geschlossen ist.

Sei ν eine Grundsubstitution, die ein Spezialisierung von σ und damit auch von τ ist. Dann gilt aufgrund von Lemma 4.7.11 für $1 \leq i \leq k$ und $1 \leq j \leq p_i$

$$(r^{ij}\nu)_\nu \in \langle t^*_{ij} \rangle_{A^n_i} \quad \text{und} \quad (\overline{r}^{ij}\nu)_\nu \in \langle s^*_{ij} \rangle_{A^n_i}.$$

Daraus folgt

$$r^{ij}\nu \in [t^*_{ij}\nu]^*_{(A^n_i,\nu)} \quad \text{und} \quad \overline{r}^{ij}\nu \in [s^*_{ij}\nu]^*_{(A^n_i,\nu)}$$

und damit

$$[t^*_{ij}\nu]^*_{(A^n_i,\nu)} = [s^*_{ij}\nu]^*_{(A^n_i,\nu)}.$$

Da die Methode $\mathcal{M}_0$ verwendet wird, mit der keine universellen Formeln gefunden werden, ist $[t^*_{ij}\nu]^*_{(A^n_i,\nu)} = [t^*_{ij}\nu]^*_{(A^n_i\nu,\nu)}$ bzw. $[s^*_{ij}\nu]^*_{(A^n_i,\nu)} = [s^*_{ij}\nu]^*_{(A^n_i\nu,\nu)}$, und daher

$$[(t^*_{ij}\nu)\nu]^*_{(A^n_i\nu,\nu)} = [(s^*_{ij}\nu)\nu]^*_{(A^n_i\nu,\nu)}.$$

Dann ist aber das Tableau $T_n\nu$ nach Definition 4.3.4 und mit Satz 4.3.7 auch nach Definition 4.7.10 geschlossen. Schließlich gilt $T_n\tau\nu = T_n\nu$, denn τ ist kanonisch und allgemeiner als ν. Also ist auch $T_n\tau$ nach Definition 4.7.10 geschlossen. ■

4.9 Mögliche Änderungen an der Berechnung von $\langle t \rangle_A$

4.9.1 Anwendung von Gleichungen aufeinander

Bei dem im vorigen Abschnitt beschriebenen Verfahren zur Berechnung von $\langle t \rangle_A$ werden zum Abschluß eines Astes keine Gleichungen auf Gleichungen angewendet. Dadurch wird zwar die Vollständigkeit nicht eingeschränkt, die Anwendung von Gleichungen auf Gleichungen könnte aber den Vorteil haben, daß Gleichungsketten, die zum Abschluß eines Astes mehrfach angewendet werden müssen, zu einer Gleichung zusammengefaßt werden.

Man bräuchte dazu aber eine gute Heuristik, die festlegt, welche Gleichungsketten zusammengefaßt werden sollen. Fehlt sie, und werden Gleichungen wahllos auf Gleichungen angewendet, um neue Gleichungen zu erzeugen, ist dies zumeist von Nachteil — insbesondere dann, wenn es keine Gleichungsketten gibt, die mehrfach angewendet werden, oder diese Ketten relativ kurz sind. Auf die Anwendungen von Gleichungen auf Gleichungen wird darum verzichtet.

Das Fehlen einer guten Heuristik steht auch im Zusammenhang mit der Nichtanwendbarkeit des Knuth–Bendix–Verfahrens (Abschnitt 2.4).

4.9.2 Hinzufügen nur jeweils eines Terms

Zur Berechnung von $\langle t \rangle^{n+1}_A$ aus $\langle t \rangle^n_A$ werden nach Definition 4.7.4 alle aus $\mathcal{H}(\langle t \rangle^n_A)$ ableitbaren Terme betrachtet. Eine andere Möglichkeit wäre, nur jeweils ein neues Element hinzuzufügen, das durch eine weitere Heuristik ausgewählt würde.

Man hätte dann eine weitere Möglichkeit an der Hand, das Wachstum der Mengen $\langle t \rangle^n_A$ zu begrenzen und in eine günstige, durch eine Heuristik bestimmte Richtung zu lenken. Denkbar wäre insbesondere, zunächst nur solche Terme hinzuzufügen, die ein geringeres Gewicht haben, als der Term, aus dem sie abgeleitet sind.

Bei genauerer Überlegung stellt man aber fest, daß dies nicht wirklich von Vorteil ist. Will man nämlich einen bestimmten, besonders günstigen der aus $t_n = \mathcal{H}(\langle t \rangle_A^n)$ ableitbaren Terme auswählen, muß man zunächst *alle* diese ableitbaren Terme bestimmen. Dann kann man sie aber auch gleich zu $\langle t \rangle_A^{n+1}$ hinzufügen, und es der Auswahlheuristik $\mathcal{H}$ überlassen, die ungünstigeren dieser Terme zunächst nicht weiter zu betrachten.

Allenfalls dann wäre es besser, zunächst nur einen Term hinzuzufügen, wenn man über eine Möglichkeit verfügte, schon den auf den Term t_n anwendbaren Gleichungen „anzusehen“, welche von ihnen zu günstigeren Termen führen. Dies ist in der Regel aber schwierig und würde die Freiheiten, die man bei der Definition des Termgewichtes und einer darauf aufbauenden Auswahlheuristik hat, stark einschränken.

So kann man beispielsweise daraus, daß eine Seite einer Gleichung ein geringeres Gewicht hat als die andere, nicht schließen, daß die Anwendung in Richtung des geringeren Gewichtes das Gewicht eines Terms verringert. Es könnten nämlich, um die Anwendung zu ermöglich, Variablenbelegungen notwendig werden.

Beispiel 4.9.1

Wählt man als Termgewicht die Länge eines Terms und wendet die Gleichung $f(x, x) \approx x$ von links nach rechts auf den Term $h(f(g(a, b), y), y, y)_{id}$ an, so erhält man den neuen Term $h(g(a, b), g(a, b), g(a, b)_{\{x \leftarrow g(a,b), y \leftarrow g(a,b)\}}$.

Obwohl man also die Gleichung $f(x, x) \approx x$ in Richtung des geringeren Gewichtes, von links nach rechts, angewendet hat, ist ein Term mit größerem Gewicht entstanden. Dies zeigt, daß es nicht möglich ist, einer Gleichung einfach „anzusehen“, ob ihre Anwendung günstig ist.

Kapitel 5

Implementierung des Verfahrens

5.1 Datenstrukturen

Alle Datenstrukturen und die Prädikate, die auf deren interne Darstellung zugreifen, sind in dem Modul `EQDATASTRUCTURES` zusammengefaßt. Die wichtigsten Datenstrukturen und ihre Komponenten sind im folgenden aufgelistet:

Die Datenstruktur `sterm` (Tabelle 5.1) stellt ein Element s_σ aus einer Menge $\langle t \rangle_A^n$ dar. Neben s_σ selbst enhält sie alle Informationen, die für die Auswahl von Termen durch die implementierte Heuristik relevant sind.

Die Datenstruktur `inequality` (Tabelle 5.2) enthält nicht die beiden Terme einer Ungleichung $t^* \not\approx s^*$, sondern die jeweils aktuellen Mengen $\langle t^* \rangle_{A_i}^n$ und $\langle s^* \rangle_{A_i}^m$ (n und m brauchen nicht übereinzustimmen).

Die Datenstruktur `disjunction`, die eine Disjunktion von Ungleichungen aus $\mathrm{Dis}^*(A_i)$ modelliert, und die Datenstruktur `equality`, die eine Gleichung $(t \approx s) \in \mathrm{Gl}^*(A_i, id)$ darstellt, sind in den Tabellen 5.3 und 5.4 beschrieben.

5.2 Steuerung des Abschlusses eines Astes

Wie schon in Abschnitt 4.8 beschrieben, werden die Äste eines Tableaus einzeln nacheinander betrachtet, um $_3T^AP$s globale Steuerung der Suche nach einem Abschluß beibehalten zu können.

Es wird dafür vom Modul `equality` ein Prädikat

```
close_branch_with_equality(Branch)
```

zur Verfügung gestellt, das eine Implementierung des Verfahrens aus Definition 4.8.3 darstellt. Wird es mit dem Ast eines Tableaus als Argument aufgerufen, versucht es diesen zu schließen und führt die dazu notwendige Substitution aus. Tritt Backtracking auf, werden weitere Substitutionen gesucht und ausgeführt, die den Abschluß zulassen. `close_branch_with_equality` enthält zu diesem Zweck einen *choice point*. Das Prädikat wird aufgerufen, wenn $_3T^AP$ den Ast eines Tableaus voll expandiert hat, und es nicht gelingt, ihn ohne den Einsatz von Gleichheit, also durch komplementäre Atome, zu schließen.

Ist das Prädikat `close_branch_with_equality` aufgerufen und ein Ast A_i vorgelegt, werden zunächst die Mengen $\mathrm{Gl}^*(A_i, id)$ und $\mathrm{Dis}^*(A_i)$ bestimmt (beschrieben in Abschnitt 5.3). Wird schon dabei eine Substitution gefunden, die einen Abschluß ermöglicht, wird diese ausgeführt. Andernfalls wird mit der iterativen Berechnung der Mengen $\langle t \rangle_{A_i}$ begonnen, um einen Abschluß des Astes aufzufinden.

sterm	
term	Der Term s.
inst	Die Substitution σ.
univ_vars	Die Menge derjenigen Variablen, die bei der Ableitung eines Terms neu eingeführt wurden. Bezüglich dieser Variablen kann der Term als „universell“ angesehen werden. Sie brauchen darum niemals instantiiert zu werden.
parent	Der Term, aus dem s_σ abgeleitet wurde.
derive	Die Gleichung, mit deren Hilfe s_σ aus parent abgeleitet wurde.
number	Die Nummer des Elementes (an dieser ist auch ablesbar, zu welcher Ungleichung welcher Disjunktion es gehört).
first_eq	Die Nummer der ersten auf s_σ anwendbaren Gleichung. Ist keine Gleichung anwendbar, oder wurde s_σ schon ausgewählt, erhält first_eq den Wert 0.
weight	Das Gewicht des Terms s.
diff	Die Differenz des Gewichts von s zu dem Gewicht von parent (positiv, wenn s leichter ist).
depth	Die Anzahl der Ableitungsschritte, die notwendig waren, um s_σ aus t_{id} abzuleiten.

Tabelle 5.1: Die Datenstruktur sterm.

inequality	
left_side, right_side	Die Mengen $\langle t^*_{jl} \rangle^n_{A_i}$ und $\langle s^*_{jl} \rangle^m_{A_i}$ zu den beiden Seiten der Ungleichung $t^*_{jl} \not\approx s^*_{jl}$.
closings	Die Liste Γ_{jl} der bisher gefundenen, die Widerlegung der Ungleichung erlaubenden Substitutionen.
number	Die Nummer (j, l) der Ungleichung.
left_counter, right_counter	Die Anzahl der zu left_side und right_side schon hinzugefügten Terme. Damit können eindeutige Bezeichner für die Terme vergeben werden.

Tabelle 5.2: Die Datenstruktur inequality.

disjunction	
expdbl_ineq	Die Liste der noch erweiterbaren Ungleichungen der Disjunktion. Eine Ungleichung ist erweiterbar, wenn sie wenigstens noch ein Element t_σ enthält, auf das eine Gleichung anwendbar ist.
inexpdbl_ineq	Die Liste der nicht erweiterbaren Ungleichungen der Disjunktion.
number	Die Nummer der Disjunktion.

Tabelle 5.3: Die Datenstruktur disjunction.

equality	
left_side	Die linke Seite t der Gleichung.
right_side	Die rechte Seite s der Gleichung.
universal_vars	Liste der Variablen in t und s, die allquantifiziert sind, bezüglich derer die Gleichung also universell ist.
number	Die Nummer der Gleichung.

Tabelle 5.4: Die Datenstruktur `equality`.

5.3 Bestimmung von $\mathrm{Dis}^*(A)$ und $\mathrm{Gl}^*(A, id)$

Zur Berechnung der Mengen $\mathrm{Dis}^*(A_i)$ und $\mathrm{Gl}^*(A_i, id)$ dient das Prädikat

`extract_disjunctions_and_equalities`.

Zu den in $\mathrm{Dis}^*(A_i)$ auftretenden Termen werden die Mengen $\langle t^* \rangle^0_{A_i}$ bestimmt.

Die zusätzlichen Informationen, die in den Datenstrukturen enthalten sind, werden hinzugefügt. So werden die Gleichungen, Disjunktionen und Ungleichungen durchnumeriert, und zu den in $\langle t^* \rangle^0_{A_i}$ liegenden Termen t^*_{id} werden unter anderem die Nummern der ersten auf sie anwendbaren Gleichungen bestimmt.

Die Symmetrie der Gleichheit wird dadurch modelliert, daß für jede Gleichung $\mathsf{T}\,(t \approx s)$ in A_i, die universell ist bezüglich $x_1, \ldots, x_n$, neben $(\forall x_1) \ldots (\forall x_n)(t \approx s)$ auch $(\forall x_1) \ldots (\forall x_n)(s \approx t)$ zu $\mathrm{Gl}^*(A_i, id)$ hinzugefügt wird. Daher ist es möglich, im weiteren Gleichungen nur noch von links nach rechts anzuwenden, was die Implementierung vereinfacht. Auch wird die Ausgabe von Beweisen (Abschnitt 5.6) übersichtlicher, da die beiden Versionen einer Gleichung verschiedene Bezeichnungen erhalten.

Außer den Disjunktionen wird von `extract_disjunctions_and_equalities` auch eine Liste der Substitutionen zurückgegeben, die, ohne daß Ableitungen ausgeführt werden, einen Abschluß des Astes erlauben. Solche Substitutionen σ gibt es dann, wenn der Ast Ungleichungen $\mathsf{F}\,(t \approx s)$ enthält, und σ ein MGU von t^* und s^* ist.

5.4 Ausführung eines Iterationsschrittes

Ein Iterationsschritt wird durch das Prädikat

`equality_application_for_closing`

ausgeführt.

Zunächst muß derjenige Term t^* bestimmt werden, für den $\langle t^* \rangle^{n+1}_{A_i}$ aus $\langle t^* \rangle^n_{A_i}$ berechnet werden soll. Diese Auswahl geschieht mit Hilfe der Prädikate

`get_best_disjunction`,
`get_best_inequality_of_disjunction` und
`get_best_side_of_inequality`.

Sie modellieren jeweils eine Queue: das erste Element der Liste wird ausgewählt und nach seiner Erweiterung am Ende der Liste eingereiht. Dabei wird allerdings darauf geachtet, daß die ausgewählte Menge $\langle t^* \rangle^n_{A_i}$ erweiterbar ist, also zumindest einen Term enthält, auf den eine Gleichung anwendbar ist. Bessere Heuristiken wären denkbar, worauf in Abschnitt 6.3 näher eingegangen wird.

Das Prädikat `get_best_sterm_of_ineq_side` modelliert die in Abschnitt 5.5 beschriebene faire Auswahlheuristik $\mathcal{H}$. Nachdem mit seiner Hilfe ein Term $t'_\sigma \in \langle t^* \rangle^n_{A_i}$ ausgewählt ist, werden entsprechend Definition 4.7.4 die Menge Θ'_n und dann $\langle t^* \rangle^{n+1}_{A_i}$ bestimmt.

Es wird überprüft, ob mit den neu vorhandenen Elementen aus Θ'_n die Ungleichung, und wenn dies möglich ist, die Disjunktion und damit der Ast geschlossen werden kann. Dazu dient das Prädikat `find_closure_or_make_next_step`, das auch den *choice point* für das Backtracking enthält, das auftritt, wenn später weitere den Ast abschließende Substitutionen gesucht werden müssen.

Ist die ausgewählte Ungleichung nach Ausführung des Iterationsschrittes nicht mehr erweiterbar, und wurde bisher keine diese Ungleichung abschließende Substitution gefunden, so wird dies auch in Zukunft nicht geschehen. Darum wird die gesamte Disjunktion aus der Liste der zu betrachtenden Disjunktionen entfernt. Bleibt dannach keine Disjunktion mehr übrig, so kann der Ast nicht geschlossen werden und `close_branch_with_equality` scheitert.

5.5 Heuristik zur Auswahl eines Termes

Eine nach Definition 4.7.7 faire Heuristik $\mathcal{H}$ zur Auswahl des nächsten zu betrachtenden Elementes von $\langle t^* \rangle^n_{A_i}$ ist durch das Prädikat

```
get_best_sterm_of_ineq_side
```

gegeben. Es stützt sich darauf, daß die $\langle t \rangle^n_{A_i}$ modellierende Liste sortiert ist bezüglich einer Ordnungsrelation `is_better_term`, und wählt jeweils das erste Element der Liste aus. Neu berechnete Elemente werden entsprechend der Relation in $\langle t^* \rangle^n_{A_i}$ eingereiht.

Die Relation `is_better_term` ist durch folgende Kriterien definiert, die in der Reihenfolge ihrer Wichtigkeit genannt sind:

1. Elemente, die schon einmal ausgewählt wurden, oder auf die keine Gleichung anwendbar ist, haben die niedrigste Priorität und werden niemals ausgewählt. Für wenigstens ein Element in $\langle t^* \rangle^n_{A_i}$ trifft dies nicht zu, da Mengen $\langle t^* \rangle^n_{A_i}$, die nicht erweiterbar sind, der Auswahlheuristik gar nicht erst vorgelegt werden.

2. Das zweitwichtigste Kriterium für die Auswahl ist das Gewicht der abgeleiteten Terme. Es soll möglichst gering sein. Um diesem Ziel gerecht zu werden, wird — wie schon in Abschnitt 4.9.2 beschrieben — nicht nur das Gewicht $G(s)$ des auszuwählenden Termes s, sondern auch die Differenz $D(s)$ seines Gewichtes zum Gewicht des Termes betrachtet, aus dem er abgeleitet wurde. Bevorzugt werden in dieser Reihenfolge:

   (a) Terme s für die $D(s)$ positiv ist,

   (b) Terme s mit geringerem Gewicht $G(s)$,

   (c) Terme s mit größerer Gewichtsdifferenz $D(s)$, was sowohl eine größere Verbesserung, als auch, wenn $D(s)$ negativ ist, eine geringere Verschlechterung bedeuten kann.

   Für das Element t_{id} aus $\langle t^* \rangle^n_{A_i}$ ist $D(t)$ zwar nicht definiert, als zu Anfang einziges Element wird es aber stets ausgewählt und später nicht mehr betrachtet.

3. Das letzte Kriterium ist schließlich die Ableitungstiefe. Elemente, die durch weniger Ableitungsschritte erzeugt wurden, werden bevorzugt.

Stimmen Elemente von $\langle t^* \rangle^n_{A_i}$ bezüglich all dieser Kriterien überein, so wird dasjenige bevorzugt, das zuerst abgeleitet wurde, was dadurch sichergestellt wird, daß neue Elemente in $\langle t^* \rangle^n_{A_i}$ hinter allen gleichwertigen schon vorhandenen Elementen eingereiht werden.

Daß der Gewichtsdifferenz $G(s)$ ein so hoher Stellenwert eingeräumt wird, ist die Konsequenz aus dem in Abschnitt 4.9.2 dargelegten. Auch ungünstige Elemente s_σ, die durch eine das Gewicht $G(s)$ vergrößernde Ableitung entstehen, für die $D(s)$ also negativ ist, werden zu $\langle t^* \rangle^n_A$ hinzugefügt. Ihre weitere Betrachtung wird erst durch die Auswahlheuristik vermieden.

Das Termgewicht ist durch ein Prädikat `term_weight` gegeben. Es ist standardmäßig als die lexikographische Länge des Termes definiert, und erfüllt daher die Bedingungen aus Satz 4.7.8:

Definition 5.5.1 (Gewicht eines Termes)

Das Gewicht $G(t)$ eines Termes t ist gegeben durch

$$\begin{aligned} G(c) &= 1 \\ G(x) &= 1 \\ G(f(t_1,\ldots,t_k)) &= 1 + G(t_1) + \cdots + G(t_k) \end{aligned}$$

Dabei bezeichnet c eine Konstante, x eine Variable und f ein beliebiges k–stelliges Funktionssymbol ($k \geq 1$).

Die Definition des Termgewichtes kann aber durch den Benutzer leicht durch zusätzliche Klauseln wie

```
term_weight(e,5).
```

verändert werden. So ist es beispielsweise zum Beweis gruppentheoretischer Sätze günstig, das Gewicht des neutralen Elementes wie gezeigt zu erhöhen.

Satz 5.5.2

Erfüllt die durch `term_weight` *gegebene Gewichtsfunktion die Bedingungen aus Satz 4.7.8, ist also die Zahl der Terme mit einem bestimmten Gewicht endlich und G monoton bezüglich der Subsumtion, dann modelliert das Prädikat* `get_best_term_of_ineq_side`*, wie es oben beschrieben ist, eine faire Auswahlheuristik.*

Beweis:

Sei $s_\sigma \in \langle t^* \rangle_A^n$ beliebig und nicht schon für ein $m < n$ ausgewählt worden (sonst ist die Bedingung $\mathcal{H}(\langle t^* \rangle_A^m) = s_\sigma$ trivialerweise erfüllt.

Im Beweis des Satzes 4.7.8 wurde schon gezeigt, daß es für jedes $m \geq n$ in $\langle t^* \rangle_A^m$ ein ${s'}_{\sigma'}$ gibt, das s_σ subsumiert und für das $G(s') \leq G(s)$ gilt.

Zunächst ist festzuhalten, daß immer nur aus endlich vielen aufeinanderfolgenden Mengen $\langle t^* \rangle_A^j, \langle t \rangle_A^{j+1}, \ldots$ Terme t'_τ mit positivem $D(t')$ ausgewählt werden, denn ihr Gewicht $G(t')$ ist durch das maximale Termgewicht in $\langle t^* \rangle_A^j$ nach oben beschränkt, und damit ihre Anzahl.

Es werden also auch immer wieder Terme t'_τ mit $D(t') \leq 0$ ausgewählt. Zumindest diese Terme müssen — solange keines der ${s'}_{\sigma'}$ gewählt wird — ein Gewicht $G(t') \leq G(s') \leq G(s)$ haben. Wieder greift nun das Argument, daß die Anzahl solcher Terme endlich ist, sie also früher oder später alle ausgewählt werden. Dann muß aber wegen $G(s') \leq G(s)$ auch eines der ${s'}_{\sigma'}$ an die Reihe kommen. ■

5.6 Ausgabe der Ableitungen

Ist die globale $_3T^AP$–Variable `eqdebuglevel` von Null verschieden, wird ein Protokoll über die Suche nach dem Abschluß eines Astes mit Hilfe der Gleichheit ausgegeben. Durch den der Variablen zugewiesenen Wert (1 bis 5), kann festgelegt werden, welche Einzelheiten ausgegeben werden sollen.

Jedesmal, wenn `close_branch_with_equality` zum Abschluß eines nicht schon vorher betrachteten Astes aufgerufen wird, erhält dieser Ast eine eindeutige Bezeichnung (`b1`, `b2`, `b3`, ...).

In der Ausgabe werden alle freien Variablen durch die PROLOG–Variablen bezeichnet, durch die sie dargestellt sind, also in der Form $\_n$.

Eine Substitution $\sigma = \{\_n_1 \leftarrow t_1, \ldots, \_n_k \leftarrow t_k\}$ wird in der Form

[$\_n_1$=t_1, ..., $\_n_k$=t_k]

dargestellt, die leere Substitution id also durch `[]`.

Disjunktionen werden in der Form (n) durch ihre Nummer n bezeichnet, die k–te Ungleichung $t_{nk} \not\approx s_{nk}$ der Disjunktion n durch (n,k), das j–te Element in der zur rechten Seite der Ungleichung gehörenden Menge $\langle t^*_{nk}\rangle_{A_i}$ mit (n,k,r,j) und das j–te Element in $\langle s^*_{nk}\rangle_{A_i}$ mit (n,k,l,j).

Ein Element s_σ in einer Menge $\langle t^*\rangle_A$ wird in der Form

$\langle$*Bezeichner*$\rangle$ ($\langle$*Differenz*$\rangle$,$\langle$*Gewicht*$\rangle$) $\langle$*Term*$\rangle$ $\langle$*Substitution*$\rangle$

ausgegeben. Dabei ist $\langle$*Gewicht*$\rangle$ das Gewicht $G(s)$ des Terms und $\langle$*Differenz*$\rangle$ die Differenz $D(s)$ von $G(s)$ zu dem Gewicht des Termes aus dem s_σ abgeleitet wurde. Die beiden wichtigsten Kriterien für die Auswahl eines Termes werden also mit den Termen ausgegeben.

Ein Protokoll beginnt mit der Ausgabe der Zeile

```
Beginning search for closure of branch b1
```

Anschließend werden die Gleichungen in $\mathrm{Gl}^*(A_i, id)$ ausgegeben. Die n–te Gleichung

$$(\forall u_1)\dots(\forall u_j)(t \approx s),$$

wird dargestellt durch:

n: [_u_1,...,_u_j] t = s

Nach den Gleichungen folgen die Disjunktionen in $\mathrm{Dis}^*(A_i)$. Ist beispielsweise (f(_13) $\not\approx$ g(_13) $\vee$ $a \not\approx$_14) die einzige extrahierte Disjunktion, dann wird folgendes ausgegeben:

```
(1,1): f(_13) \= g(_13)
(1,1,r,0): (0,2) f(_13) []
(1,1,l,0): (0,2) g(_13) []

(1,2): a \= _14
(1,2,r,0): (0,1) a []
(1,2,l,0): (0,1) _14 []
```

Anschließend beginnt die iterative Berechnung der Mengen $\langle t^*\rangle^n_{A_i}$. Während dies geschieht, wird für jeden neu abgeleiten Term eine Protokollzeile ausgegeben. Dabei bedeutet beispielsweise

```
(1,1,r,0): (0,2) f(_13) []  -2->  (1,1,r,1): (1,1) f(c) [_13=a]
```

daß das neue Element $f(c)_{\{\_13\leftarrow a\}}$ mit der Bezeichnung (1,1,r,1) aus dem Element (1,1,r,0) mit Hilfe der Gleichung Nummer 2 abgeleitet und zu $\langle t^*_{1,1}\rangle_{A_i}$ hinzugefügt wurde.

Eine Zeile wie

```
Ineq. (1,1) closed by f(c) = f(c) [_13=a]
```

bedeutet, daß die Ungleichung $t_{1,1} \not\approx s_{1,1}$ mit Hilfe der Substitution $\sigma = \{\_13 \leftarrow a\}$ widerlegt werden kann, es also Elemente $f(c)_\tau \in \langle t^*_{1,1}\rangle^n_{A_i}$ und $f(c)_{\tau'} \in \langle s^*_{1,1}\rangle^{n'}_{A_i}$ gibt, und σ eine Vereinigung von τ und τ' ist.

Ist damit beispielsweise eine Substitution $\{\_13 \leftarrow a, \_14 \leftarrow a\}$ gefunden, die es erlaubt, die gesamte Disjunktion Nummer 1 zu widerlegen, und damit den Ast A_i zu schließen, wird

```
Disj. (1) closed with [_13=a,_14=b]
```

ausgegeben.

Kann ein Ast A_i direkt, ohne die Ableitung neuer Terme geschlossen werden, da er eine Ungleichung $\mathsf{F}\,(t \approx s)$ enthält, und etwa $\sigma = \{\texttt{\_13} \leftarrow a\}$ ein MGU von t und s ist, wird dies durch

```
Branch closed by one of its inequalities with [_13=a]
```

angezeigt.

Ist eine der Disjunktionen oder eine der Ungleichungen nicht mehr erweiterbar, wird

```
Disj. (3) is exhausted
```

bzw.

```
Ineq. (3,1) is exhausted
```

ausgegeben.

Wenn keine der Disjunktionen mehr erweiterbar ist, der Ast also nicht mit der Hilfe von Gleichheit geschlossen werden kann, wird dies gemeldet:

```
----- Branch is exhausted and cannot be closed with equality -----
```

Tritt dann Backtracking auf, und wird ein schon geschlossener Ast erneut betrachtet, beginnt dies mit der Ausgabe der Zeile

```
Resuming search for closure with equality of branch b17
```

Alle der Ausgabe von Ableitungen dienenden Prädikate sind im Modul `EQOUTPUT` zusammengefaßt, so daß die Form der Ausgabe leicht geändert werden kann.

Beispiele für Ausgaben des Gleichheitsmoduls sind im Anhang in Abschnitt A enthalten.

Kapitel 6

Fazit und Ausblick

6.1 Leistungsfähigkeit der Implementierung

Das implementierte Gleichheitsmodul fügt sich hinsichtlich seiner Leistungsfähigkeit im großen und ganzen gut in den Beweiser ${}_3T^AP$ ein.

So stellen Aufgaben aus Pelletiers Beispielsammlung ([Pelletier, 1986]), die dort mit einer bestimmten Punktzahl für ihren Schwierigkeitsgrad versehen sind, unabhängig davon, ob sie mit oder ohne Gleichheit formuliert sind, ${}_3T^AP$ nun in etwa vor die gleichen Probleme.

Es zeigt sich, daß für fast alle Theoreme, für die überhaupt ein Beweis gefunden wird, dies in sehr viel kürzerer Zeit geschieht, als es etwa dem Menschen möglich wäre.

Vergrößert man die Komplexität der Probleme, nimmt die benötigte Zeit zunächst nur wenig, dann aber, wenn ein gewisser Schwierigkeitsgrad erreicht ist, sehr schnell zu.

Alle hier gemachten Aussagen sind natürlich nur im Allgemeinen und im Durchschnitt gültig, denn es lassen sich sowohl spezielle Probleme angeben, die besonders auf das implementierte Verfahren zugeschnitten sind, als auch solche die besonders ungeeignet sind.

6.2 Die Verwendung der Sprache PROLOG

PROLOG eignet sich prinzipiell sehr gut für die Implementierung eines Tableaubeweisers mit freien Variablen. So können beispielsweise die freien Variablen direkt durch PROLOG–Variablen dargestellt werden. Auch ist Prolog gut geeignet zur Behandlung von Termen.

Diese Vorteile werden allerdings relativiert, wenn ein nicht–trivialer Tableaukalkül, wie der des Beweisers ${}_3T^AP$, implementiert wird. Das PROLOG–interne Backtracking kann nur sehr begrenzt eingesetzt werden, wenn eine Breitensuche modelliert werden soll.

Die freien Variablen können zwar als PROLOG–Variablen dargestellt werden, die einfache Instantiierung von PROLOG–Variablen kann jedoch nicht verwendet werden. Auch ist die Verwaltung von Daten, die global zur Verfügung stehen sollen, oft schwierig.

Insgesamt stellt sich also die Frage, ob für die Implementierung von ${}_3T^AP$ und der Gleichheit eine imperative oder applikative Sprache nicht besser geeignet gewesen wäre.

6.3 Mögliche Erweiterungen

Im folgenden sind einige mögliche Erweiterungen des Gleichheitsmoduls zusammengestellt, von denen die ersten beiden nach Abschluß dieser Studienarbeit verwirklicht werden sollen bzw. schon verwirklicht worden sind:

Demodulatoren Dem Benutzer des Beweisers soll die Möglichkeit gegeben werden, Gleichungen als Demodulatoren zu kennzeichnen.

Solche Gleichungen werden dann

- nur noch in eine Richtung, von links nach rechts, angewendet, und
- auf jeden auftretenden Term, auf den sie anwendbar sind, sofort und so oft wie möglich angewendet.

Es bleibt dem Benutzer überlassen, darauf zu achten, daß durch die Deklaration von Demodulatoren die Vollständigkeit nicht verloren geht, was im allgemeinen der Fall ist.

Durch den geschickten Einsatz von Demodulatoren kann in vielen Fällen die Ableitungskette zum Beweis der Gleichheit zweier Terme verkürzt und der Suchraum stark eingeschränkt werden.

Behandlung von Sorten In einer neuen, inzwischen fertig gestellten Version von $_3T^AP$ können Terme und Variablen mit Sorten versehen werden.

Würde auch das Gleichheitsmodul um die Verwendung von Sorten–Informationen erweitert, könnte dadurch in gewissen Fällen die Suche nach einem Abschluß stark beschleunigt werden, da die Zahl der auf bestimmte Terme anwendbaren Gleichungen eingeschränkt würde, wie auch die Zahl der potentiell einen Ast abschließenden Atome.

Knuth–Bendix–Verfahren Unter gewissen Einschränkungen (siehe Abschnitt 2.4) ist auch das Knuth–Bendix–Verfahren für den Tableaukalkül geeignet. Zum einen natürlich dann, wenn tatsächlich alle Gleichungen auf einem Ast universell sind bezüglich der Variablen, die sie enthalten, was in praktischen Anwendungen relativ häufig der Fall ist. Aber auch in den Fällen, in denen sich nicht–universelle Gleichungen auf einem Ast befinden, könnte man das Knuth–Benix–Verfahren verwenden, um die durch die universellen Gleichungen gebildete Theorie zu vervollständigen. Anschließend müßte dann aber ein anderes Verfahren eingesetzt werden, um zu untersuchen, in welcher Weise die nicht–universellen Gleichungen zu einem Abschluß beitragen können, und welche Variablensubstitutionen dafür notwendig wären. Es würde also ein möglichst großer Teil der Suche nach einem Abschluß dem effizienten Knuth–Bendix–Verfahren überlassen.

Eine weitere Einsatzmöglichkeit für das Knuth–Bendix–Verfahren wäre, zunächst die Annahme zu machen, daß alle Gleichungen auf einem Ast universell sind, und alle diejenigen Paare von Literalen zu suchen, die unter dieser Annahme einen Abschluß bilden. Anschließend müßte dann aber wiederum ein anderes Verfahren eingesetzt werden, um zu überprüfen, welche der gefundenen Paare tatsächlich einen Abschluß bildet, und welche Variablensubstitutionen notwendig sind. Man könnte so ausnutzen, daß das Knuth–Bendix–Verfahren, wenn es im Tableaukalkül eingesetzt wird, zwar nicht korrekt aber doch vollständig ist.

Eine geschickte auf dem Knuth–Bendix–Verfahren beruhende Implementierung der Gleichheitsbehandlung würde die Mächtigkeit und Effizienz des Beweisers vermutlich beträchtlich steigern und sollte auf jeden Fall versucht werden.

Weitere Auswahlheuristiken Andere Heuristiken $\mathcal{H}$ für die Auswahl von Elementen aus $\langle t^* \rangle_A^n$ sind denkbar, die auf bestimmte Problemstellungen und Anwendungsgebiete zugeschnitten sind.

Möglich wäre auch, die mit einem Element $s_\sigma \in \langle t^* \rangle_{A_i}^n$ verbundene Substitution σ in die Berechnung des Termgewichtes einzubeziehen, was bisher nicht geschieht.

Auch könnte die Auswahl der nächsten zu behandelnden Disjunktion und der nächsten Ungleichung verbessert werden. Beispielsweise dadurch, daß Disjunktionen, die nur wenige Ungleichungen enthalten, und darum eher abgeschlossen werden können, in gewissem Umfang bevorzugt werden.

Allgemeine Breitensuche Eine vielversprechende Verbesserung wäre möglicherweise auch, die Äste des Tableaus nicht der Reihe nach zu betrachten, sondern tatsächlich eine Breitensuche nach einer abschließenden Substitution durchzuführen — also das in Definition 4.8.1 beschriebene Verfahren zu implementieren.

Dafür sprächen die schon in Abschnitt 4.5 dargelegten Vorteile einer Breitensuche.

Auch müßte — wenn die Gleichheitsbehandlung ausgeschaltet ist — die Anwendung der γ–Regel nicht mehr begrenzt werden. Die Problematik der Wahl der Grenze q entfiele. Genauso entfiele die Begrenzung der Ableitungslänge bei der Anwendung von Gleichungen, womit auch die Vollständigkeit des Beweisers nicht mehr eingeschränkt wäre.

Einziger Nachteil wäre der sehr viel größere Speicheraufwand, da nicht mehr nur ein Ast, sonder das ganze Tableau zugleich zur Verfügung stehen müßte. Dieser Nachteil kann aber, da der zur Verfügung stehende Speicher begrenzt ist, entscheidend sein.

Anhang A

Beispiele

A.1 Darstellung der Beispiele

Im folgenden sind einige Beispiele für den Abschluß eines Tableaus mit Hilfe des um die Behandlung von Gleichheit erweiterten Beweisers $_3T^AP$ angegeben. Die meisten der Probleme stammen aus [Pelletier, 1986].

Wiedergegeben ist bei allen Beispielen nur die Suche nach dem Abschluß der schon vorher voll expandierten Äste des Tableaus, und zwar so wie es in Abschnitt 5.6 beschrieben ist.

Die Grenze q für die Anwendung der γ-Regel, mit der die Beweise geführt sind, ist die jeweils niedrigste, mit der die Beweise gelingen. Dadurch wird die Darstellung der Beweise möglichst übersichtlich gehalten.

Als maximale Länge von Ableitungsketten wurde in allen Fällen $r = 10$ gewählt. Wie sich zeigt ist dies ein günstiger Wert, der den Beweis all dieser so verschiedenen Theoreme erlaubt, obwohl er doch noch relativ klein ist, und das Backtracking, wo es auftritt, nicht zu stark behindert.

Der in der Statistik mit „Geschl. Äste" bezeichnete Wert beinhaltet auch diejenigen Abschlüsse, die wieder verworfen wurden, da Backtracking auftrat. Die Anzahl dieser wieder verworfenen Abschlüsse ist mit „Backtracking" bezeichnet.

Die angegebenen Zeiten wurden auf einer SUN SPARC SLC Workstation ermittelt. Sie beziehen sich auf den Gesamtbeweis und hängen darum nicht nur von der Implementierung des Gleichheits–Moduls, sondern auch von der der anderen Module des $_3T^AP$–Systems ab. Da sie aber eine Obergrenze für die für Gleichheitsanwendungen verbrauchten Zeiten darstellen, sind sie dennoch aussagekräftig.

A.2 Grundlegende Eigenschaften der Gleichheit

Den Anfang sollen Beweise einiger fundamentaler Eigenschaften der Gleichheit machen, nämlich diejenigen einer Äquivalenzrelation: Reflexivität, Symmetrie und Transitivität.

Anschließend wird eine einfache Anwendung des ebenfalls grundlegenden Ersetzungsaxioms bewiesen:

$$\{a \approx b,\ P(a)\} \models P(b).$$

$$\begin{array}{c}
\mathsf{F}\,(\forall x)(\forall y)(x \approx y \supset y \approx x) \\
| \\
\mathsf{F}\,(\forall y)(c \approx y \supset y \approx c) \\
| \\
\mathsf{F}\,(c \approx d \supset d \approx c) \\
| \\
\mathsf{T}\,(c \approx d) \\
| \\
\mathsf{F}\,(d \approx c) \\
\texttt{b1}
\end{array}$$

Abbildung A.1: Das voll expandierte Tableau zum Beweis der Symmetrie der Gleichheit.

A.2.1 Reflexivität der Gleichheit

Zu beweisen:

$$\models (\forall x)(x \approx x)$$

Das voll expandierte Tableau:

$$\begin{array}{c}
\mathsf{F}\,(\forall x)(x \approx x) \\
| \\
\mathsf{F}\,(c \approx c) \\
\texttt{b1}
\end{array}$$

Da der einzige Ast des Tableaus die Ungleichung $\mathsf{F}\,(c \approx c)$ enthält, sind keine Ableitungsschritte notwendig. Auch müssen keine freien Variablen belegt werden:

```
Beginning search for closure with equality of branch b1

Equalities extracted from Branch: none

Disjunctions extracted from Branch: none

Branch closed by one of its inequalities with  []
-----------------------------------------------

-------------------- PROOF --------------------
```

Geschl. Äste	Backtracking	Laufzeit	Grenze q
1	0	0,07s	1

A.2.2 Symmetrie der Gleichheit

Zu beweisen:

$$\models (\forall x)(\forall y)(x \approx y \supset y \approx x)$$

Das voll expandierte Tableau ist in Abbildung A.1 dargestellt.

Wieder enthält das Tableau nur einen Ast. Die Ungleichung $\mathsf{F}\,(d \approx c)$ kann mit Hilfe der Gleichung $\mathsf{T}\,(c \approx d)$ in einem Ableitungsschritt widerlegt werden:

$$\mathsf{F}\,(\forall x)(\forall y)(\forall z)(x \approx y \land y \approx z \supset x \approx z)$$
$$|$$
$$\mathsf{F}\,(\forall y)(\forall z)(a \approx y \land y \approx z \supset a \approx z)$$
$$|$$
$$\mathsf{F}\,(\forall z)(a \approx b \land b \approx z \supset a \approx z)$$
$$|$$
$$\mathsf{F}\,(a \approx b \land b \approx c \supset a \approx c)$$
$$|$$
$$\mathsf{T}\,(a \approx b \land b \approx c)$$
$$|$$
$$\mathsf{F}\,(a \approx c)$$
$$|$$
$$\mathsf{T}\,(a \approx b)$$
$$|$$
$$\mathsf{T}\,(b \approx c)$$

`b1`

Abbildung A.2: Das voll expandierte Tableau zum Beweis der Transitivität der Gleichheit.

```
Beginning search for closure with equality of branch b1

Equalities extracted from Branch:
1:  []  c  =  d
2:  []  d  =  c


Disjunctions extracted from Branch:
(1,1): d  \=  c  (expdbl.)

(1,1,l,0): (0,1) d []
(1,1,r,0): (0,1) c []


(1,1,l,0): (0,1) d  -2->  (1,1,l,1): (0,1) c []
Ineq. (1,1)  closed by  c  =  c  []
Disj. (1)  closed with  []
----------------------------------------------

-------------------- PROOF --------------------
```

Geschl. Äste	Backtracking	Laufzeit	Grenze q
1	0	0,10s	1

A.2.3 Transitivität der Gleichheit

Zu beweisen:

$$\models (\forall x)(\forall y)(\forall z)(x \approx y \land y \approx z \supset x \approx z)$$

Das voll expandierte Tableau ist in Abbildung A.2 dargestellt.

Das Tableau enthält wieder nur einen Ast. Um die auf ihm liegende Ungleichung zu widerlegen, sind nun aber zwei Ableitungsschritte notwendig:

```
Beginning search for closure with equality of branch b1

Equalities extracted from Branch:
1:  []  a  =  b
2:  []  b  =  a
3:  []  b  =  c
4:  []  c  =  b


Disjunctions extracted from Branch:
(1,1): a  \=  c  (expdbl.)

(1,1,l,0): (0,1) a []
(1,1,r,0): (0,1) c []


(1,1,l,0): (0,1) a  -1->  (1,1,l,1): (0,1) b []
(1,1,r,0): (0,1) c  -4->  (1,1,r,1): (0,1) b []
Ineq. (1,1)  closed by  b  =  b  []
Disj. (1)  closed with  []
----------------------------------------------

-------------------- PROOF --------------------
```

Geschl. Äste	Backtracking	Laufzeit	Grenze q
1	0	0,13s	1

A.2.4 Anwendung des Ersetzungsaxioms

Zu beweisen:

$$\{a \approx b,\, P(a)\} \models P(b)$$

Das voll expandierte Tableau:

$$\begin{array}{c} \mathsf{T}\,(a \approx b) \\ | \\ \mathsf{T}\,P(a) \\ | \\ \mathsf{F}\,P(b) \end{array}$$

b1

Dies ist das erste Beispiel für die Bildung einer Disjunktion von Ungleichungen aus einem Paar komplementärer Atome:

```
Beginning search for closure with equality of branch b1

Equalities extracted from Branch:
1:  []  a  =  b
2:  []  b  =  a

Disjunctions extracted from Branch:
(1,1): a  \=  b  (expdbl.)

(1,1,l,0): (0,1) a []
(1,1,r,0): (0,1) b []
```

```
(1,1,l,0): (0,1) a  -1->  (1,1,l,1): (0,1) b []
Ineq. (1,1)  closed by  b  =  b  []
Disj. (1)  closed with  []
----------------------------------------------

-------------------- PROOF --------------------
```

Geschl. Äste	Backtracking	Laufzeit	Grenze q
1	0	0,08s	1

A.3 Das in Kapitel 4 eingeführte Beispiel

Als nächstes Beispiel soll noch einmal das schon in Kapitel 4 verwendete und dort beschriebene Tableau mit den beiden Ästen A_{Bsp} und A'_{Bsp} aufgegriffen werden, wie es in Abbildung 4.1 (a) dargestellt ist (der linke Ast wird zuerst behandelt):

```
Beginning search for closure with equality of branch b1

Equalities extracted from Branch:
1:  []  b  =  c
2:  []  c  =  b
3:  [_1]  g(f(_1))  =  _1
4:  [_1]  _1  =  g(f(_1))


Disjunctions extracted from Branch:
(1,1): g(g(a))  \=  a  (expdbl.)
(1,2): b  \=  c  (expdbl.)

(1,2,l,0): (0,1) b []
(1,2,r,0): (0,1) c []


(1,1,l,0): (0,3) g(g(a)) []
(1,1,r,0): (0,1) a []

(2,1): a  \=  _2  (expdbl.) closed by [[_2=a]]

(2,1,r,0): (0,1) a []
(2,1,l,0): (0,1) _2 []  (not expdbl.)


Branch closed by one of its inequalities with  [_2=a]
----------------------------------------------
```

Man beachte, daß nun die Variable x_2 mit a belegt wird. Darum enthalten die aus dem rechten Ast des Tableaus, der hier mit `b2` bezeichnet wird, extrahierten Gleichungen 3 und 4 keine freien Variablen mehr:

```
Beginning search for closure with equality of branch b2

Equalities extracted from Branch:
1:  []  b  =  c
2:  []  c  =  b
3:  []  g(a)  =  f(a)
4:  []  f(a)  =  g(a)
5:  [_1]  g(f(_1))  =  _1
6:  [_1]  _1  =  g(f(_1))
```

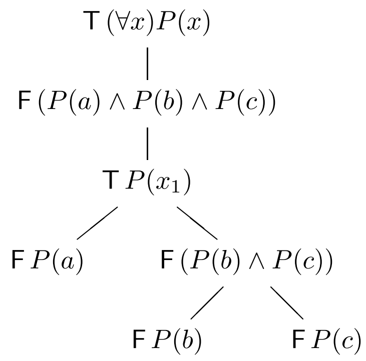


Abbildung A.3: Das für $q = 1$ voll expandierte Beispieltableau für die Verwendung universeller Formeln.

```
Disjunctions extracted from Branch:
(1,1): g(g(a))  \=  a  (expdbl.)
(1,2): b  \=  c  (expdbl.)

(1,2,l,0): (0,1) b []
(1,2,r,0): (0,1) c []


(1,1,l,0): (0,3) g(g(a)) []
(1,1,r,0): (0,1) a []


(1,1,l,0): (0,3) g(g(a))  -3->  (1,1,l,1): (0,3) g(f(a)) []
(1,1,l,0): (0,3) g(g(a))  -6->  (1,1,l,2): (-2,5) g(f(g(g(a)))) []
(1,1,l,0): (0,3) g(g(a))  -6->  (1,1,l,3): (-2,5) g(g(f(g(a)))) []
(1,1,l,0): (0,3) g(g(a))  -6->  (1,1,l,4): (-2,5) g(g(g(f(a)))) []
(1,2,l,0): (0,1) b  -1->  (1,2,l,1): (0,1) c []
(1,2,l,0): (0,1) b  -6->  (1,2,l,2): (-2,3) g(f(b)) []
Ineq. (1,2)  closed by  c  =  c  []

(1,1,r,0): (0,1) a  -6->  (1,1,r,1): (-2,3) g(f(a)) []
Ineq. (1,1)  closed by  g(f(a))  =  g(f(a))  []
Disj. (1)  closed with  []
----------------------------------------------

-------------------- PROOF --------------------
```

Geschl. Äste	Backtracking	Laufzeit	Grenze q
2	0	0,317s	1

A.4 Beispiel für die Verwendung einer universellen Formel

Das folgende Beispiel zeigt, wie der Abschluß eines Tableaus dadurch erleichtert werden kann, daß Formeln als universell erkannt werden — auch dann wenn es sich nicht um Gleichungen handelt.

Zu zeigen ist:

$$\{(\forall x)P(x)\} \models P(a) \wedge P(b) \wedge P(c)$$

Für $q = 1$ erhält man das in Abbildung A.3 dargestellte voll expandierte Tableau.

Ist es möglich, die Formel $\mathsf{T}\,P(x_1)$ als universell bezüglich x_1 zu erkennen, so kann dieses Tableau

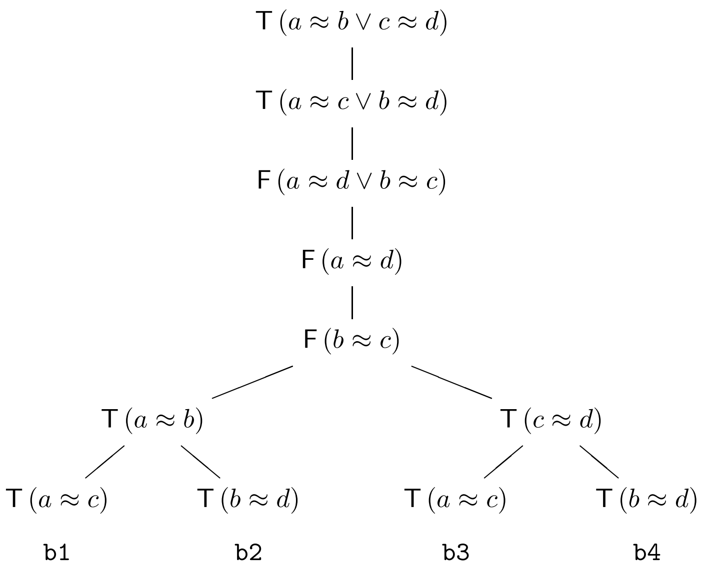


Abbildung A.4: Das voll expandierte Tableau für Pelletiers 48. Problem.

geschlossen werden. Andernfalls muß die γ–Regel mehrmals auf $\mathsf{T}\,(\forall x)P(x)$ angewendet werden, um das Tableau zu schließen.

A.5 Beispiele von Pelletier

A.5.1 Pelletiers 48. Problem

Zu zeigen:

$$\{a \approx b \vee c \approx d,\, a \approx c \vee b \approx d\} \models a \approx d \vee b \approx c$$

Das voll expandierte Tableau ist in Abbildung A.4 dargestellt.

Dies ist das erste Beispiel, bei dem das mit Hilfe der Gleichheit zu schließende Tableau mehrere Äste hat, wenn auch noch keine zueinander passenden, die Äste abschließenden Substitutionen gefunden werden müssen:

```
Beginning search for closure with equality of branch b1

Equalities extracted from Branch:
1:  []  a  =  b
2:  []  b  =  a
3:  []  a  =  c
4:  []  c  =  a

Disjunctions extracted from Branch:
(1,1): a  \=  d  (expdbl.)

(1,1,l,0): (0,1) a []
(1,1,r,0): (0,1) d []  (not expdbl.)

(2,1): b  \=  c  (expdbl.)

(2,1,l,0): (0,1) b []
(2,1,r,0): (0,1) c []
```

```
(1,1,l,0): (0,1) a  -1->  (1,1,l,1): (0,1) b []
(1,1,l,0): (0,1) a  -3->  (1,1,l,2): (0,1) c []
(2,1,l,0): (0,1) b  -2->  (2,1,l,1): (0,1) a []
(2,1,r,0): (0,1) c  -4->  (2,1,r,1): (0,1) a []
Ineq. (2,1)  closed by  a  =  a  []
DIsj. (2)  closed with  []
----------------------------------------------

Beginning search for closure with equality of branch b2

Equalities extracted from Branch:
1:  []  a  =  b
2:  []  b  =  a
3:  []  b  =  d
4:  []  d  =  b

Disjunctions extracted from Branch:
(1,1): a  \=  d  (expdbl.)

(1,1,l,0): (0,1) a []
(1,1,r,0): (0,1) d []

(2,1): b  \=  c  (expdbl.)

(2,1,l,0): (0,1) b []
(2,1,r,0): (0,1) c []  (not expdbl.)


(1,1,l,0): (0,1) a  -1->  (1,1,l,1): (0,1) b []
(2,1,l,0): (0,1) b  -2->  (2,1,l,1): (0,1) a []
(2,1,l,0): (0,1) b  -3->  (2,1,l,2): (0,1) d []
(1,1,r,0): (0,1) d  -4->  (1,1,r,1): (0,1) b []
Ineq. (1,1)  closed by  b  =  b  []
DIsj. (1)  closed with  []
----------------------------------------------

Beginning search for closure with equality of branch b3

Equalities extracted from Branch:
1:  []  b  =  d
2:  []  d  =  b
3:  []  c  =  d
4:  []  d  =  c

Disjunctions extracted from Branch:
(1,1): d  \=  a  (expdbl.)

(1,1,r,0): (0,1) d []
(1,1,l,0): (0,1) a []  (not expdbl.)

(2,1): b  \=  c  (expdbl.)

(2,1,l,0): (0,1) b []
(2,1,r,0): (0,1) c []


(1,1,r,0): (0,1) d  -2->  (1,1,r,1): (0,1) b []
(1,1,r,0): (0,1) d  -4->  (1,1,r,2): (0,1) c []
(2,1,l,0): (0,1) b  -1->  (2,1,l,1): (0,1) d []
(2,1,r,0): (0,1) c  -3->  (2,1,r,1): (0,1) d []
Ineq. (2,1)  closed by  d  =  d  []
Disj. (2)  closed with  []
----------------------------------------------
```

```
Beginning search for closure with equality of branch b4

Equalities extracted from Branch:
1:  []  a  =  c
2:  []  c  =  a
3:  []  c  =  d
4:  []  d  =  c

Disjunctions extracted from Branch:
(1,1): a  \=  d  (expdbl.)

(1,1,l,0): (0,1) a []
(1,1,r,0): (0,1) d []

(2,1): c  \=  b  (expdbl.)

(2,1,r,0): (0,1) c []
(2,1,l,0): (0,1) b []  (not expdbl.)


(1,1,l,0): (0,1) a  -1->  (1,1,l,1): (0,1) c []
(2,1,r,0): (0,1) c  -2->  (2,1,r,1): (0,1) a []
(2,1,r,0): (0,1) c  -3->  (2,1,r,2): (0,1) d []
(1,1,r,0): (0,1) d  -4->  (1,1,r,1): (0,1) c []
Ineq. (1,1)  closed by  c  =  c  []
Disj. (1)  closed with  []
----------------------------------------------

-------------------- PROOF --------------------
```

Geschl. Äste	Backtracking	Laufzeit	Grenze q
4	0	0,25s	1

A.5.2 Pelletiers 49. Problem

Zu zeigen:

$$\{(\exists x)(\exists y)(\forall z)(z \approx x \lor z \approx y),\ P(a) \land P(b),\ \neg(a \approx b)\} \models (\forall x)P(x)$$

Der Beweis gelingt erst mit einer Grenze $q \geq 3$ für die Anwendung der γ–Regel. Das in Abbildung A.5 dargestellte, für $q = 1$ voll expandierte Tableau kann nicht geschlossen werden.

Schon der Abschluß des ersten Astes `b1` gelingt nicht, wie das folgende Protokoll zeigt:

```
Beginning search for closure with equality of branch b1

Equalities extracted from Branch:
1:  []  _1  =  f2
2:  []  f2  =  _1

Disjunctions extracted from Branch:
(1,1): a  \=  f1  (expdbl.)

(1,1,l,0): (0,1) a []
(1,1,r,0): (0,1) f1 []

(2,1): b  \=  f1  (expdbl.)

(2,1,l,0): (0,1) b []
(2,1,r,0): (0,1) f1 []
```

$$\mathsf{T}\,(\exists x)(\exists y)(\forall z)(z \approx x \vee z \approx y)$$

$$\mathsf{T}\,(P(a) \wedge P(b))$$

$$\mathsf{T}\,\neg(a \approx b)$$

$$\mathsf{F}\,(\forall x)P(x)$$

$$\mathsf{F}\,P(f_1)$$

$$\mathsf{T}\,P(a)$$

$$\mathsf{T}\,P(b)$$

$$\mathsf{F}\,(a \approx b)$$

$$\mathsf{T}\,(\exists y)(\forall z)(z \approx f_2 \vee z \approx y)$$

$$\mathsf{T}\,(\forall z)(z \approx f_2 \vee z \approx f_3)$$

$$\begin{array}{cc} \mathsf{T}\,(z \approx f_2) & \mathsf{T}\,(z \approx f_3) \\ \texttt{b1} & \texttt{b2} \end{array}$$

Abbildung A.5: Für $q = 1$ voll expandiertes Tableau für Pelletiers 49. Problem.

```
(3,1): a  \=  b  (expdbl.)

(3,1,l,0): (0,1) a []
(3,1,r,0): (0,1) b []


(1,1,l,0): (0,1) a  -1->  (1,1,l,1): (0,1) f2 [_1=a]
(2,1,l,0): (0,1) b  -1->  (2,1,l,1): (0,1) f2 [_1=b]
(3,1,l,0): (0,1) a  -1->  (3,1,l,1): (0,1) f2 [_1=a]
(1,1,r,0): (0,1) f1  -1->  (1,1,r,1): (0,1) f2 [_1=f1]
(2,1,r,0): (0,1) f1  -1->  (2,1,r,1): (0,1) f2 [_1=f1]
(3,1,r,0): (0,1) b  -1->  (3,1,r,1): (0,1) f2 [_1=b]
Ineq. (1,1) is exhausted
Disj. (1) is exhausted
Ineq. (2,1) is exhausted
Disj. (2) is exhausted
Ineq. (3,1) is exhausted
Disj. (3) is exhausted
----- Branch is exhausted and cannot be closed with equality -----
```

Das voll expandierte Tableau für $q = 3$ ist zu umfangreich, um es hier wiederzugeben. Bei der Suche nach einem Abschluß dieses Tableaus tritt Backtracking auf, da die verschiedenen durch die Anwendung der γ–Regel entstehenden freien Variablen in bestimmter Weise belegt werden müssen. An den mit „$\cdots$“ gekennzeichneten Stellen sind Teile des Beweises ausgelassen.

```
Beginning search for closure with equality of branch b1

Equalities extracted from Branch:
1:  []  _5322  =  f14
2:  []  f14  =  _5322
3:  []  _6211  =  f15
4:  []  f15  =  _6211
5:  []  _6848  =  f14
6:  []  f14  =  _6848

Disjunctions extracted from Branch:
(1,1): a  \=  f13  (expdbl.)

...


Disj. (1)  closed with  [_6848=a,_5322=f13]
----------------------------------------------

Beginning search for closure with equality of branch b2

Equalities extracted from Branch:
1:  []  a  =  f15
2:  []  f15  =  a
3:  []  f13  =  f14
4:  []  f14  =  f13
5:  []  _6211  =  f15
6:  []  f15  =  _6211

Disjunctions extracted from Branch:
(1,1): a  \=  f13  (expdbl.)


...


Disj. (1)  closed with  [_6211=f13]
----------------------------------------------
```

```
Beginning search for closure with equality of branch b3

Equalities extracted from Branch:
1:  []  _12990  =  f15
2:  []  f15  =  _12990
3:  []  f13  =  f14
4:  []  f14  =  f13

Disjunctions extracted from Branch:
(1,1): a  \=  f13  (expdbl.)


...


----- Branch is exhausted and cannot be closed with equality -----

Resuming search for closure with equality of branch b2

Ineq. (3,1)  closed by  f15  =  f15  [_6211=b]
Disj. (3)  closed with  [_6211=b]
-----------------------------------------------

Beginning search for closure with equality of branch b4


...


Beginning search for closure with equality of branch b10

Equalities extracted from Branch:
1:  []  a  =  f14
2:  []  f14  =  a
3:  []  b  =  f14
4:  []  f14  =  b
5:  []  f13  =  f15
6:  []  f15  =  f13

Disjunctions extracted from Branch:
(1,1): a  \=  f13  (expdbl.)


...


Disj. (3)  closed with  []
-----------------------------------------------

-------------------- PROOF --------------------
```

Geschl. Äste	Backtracking	Laufzeit	Grenze q
10	2	1,50s	3

A.5.3 Pelletiers 55. Problem

Das folgende Problem, eine ursprünglich von Len Schubert stammende Logelei, ist ebenfalls [Pelletier, 1986] entnommen. In natürlicher Sprache lautet es:

> Jemand, der auf Dreadsbury Mansion lebt, hat Tante Agatha ermordet. Agatha, der Butler und Charles leben auf Dreadsbury Mansion, und sie sind die einzigen, die dort leben. Ein

> Mörder haßt immer sein Opfer und ist niemals reicher als sein Opfer. Charles haßt niemanden, den Tante Agatha haßt. Agatha haßt jeden außer dem Butler. Der Butler haßt jeden, der nicht reicher ist als Tante Agatha. Der Butler haßt jeden, den Agatha haßt. Niemand haßt alle. Agatha ist nicht der Butler.
>
> Daraus folgt: Agatha hat Selbstmord begangen.

Die Aufgabe kann folgendermaßen prädikatenlogisch formuliert werden:

$$\begin{array}{c}
(\exists x)(L(x) \wedge K(x,a)) \\
L(a) \wedge L(b) \wedge L(c) \\
(\forall x)(L(x) \supset (x \approx a \vee x \approx b \vee x \approx c)) \\
(\forall x)(\forall y)(K(x,y) \supset H(x,y)) \\
(\forall x)(\forall y)(K(x,y) \supset \neg R(x,y)) \\
(\forall x)(H(a,x) \supset \neg H(c,x)) \\
(\forall x)(\neg(x \approx b) \supset H(a,x)) \\
(\forall x)(\neg R(x,a) \supset H(b,x)) \\
(\forall x)(H(a,x) \supset H(b,x)) \\
(\forall x)(\exists y)\neg H(x,y) \\
\neg(a \approx b) \\
\hline
K(a,a)
\end{array}$$

Dabei haben die Prädikatensymbole die folgende Bedeutung:

$L(x)$	x lebt auf Dreadsbury Mansion.
$K(x,y)$	x ist Mörder von y.
$H(x,y)$	x haßt y.
$R(x,y)$	x ist reicher als y.

Sowohl das Tableau als auch der Nachweis seines Abschlusses sind zu umfangreich, um hier wiedergegeben zu werden. Interessant ist dieses Beispiel wegen der natürlichsprachlichen Formulierung, die es dem Menschen erlaubt, seinen Schwierigkeitsgrad zu beurteilen. Es gehört durchaus noch zu der Klasse von Problemen, die das $_3T^AP$–System besser als ein Mensch lösen kann, was bei zunehmender Komplexität von Problemen nicht mehr der Fall ist.

Für $q = 1$ werden 37 Äste geschlossen. Sechs davon mit Hilfe von Gleichheit:

Geschl. Äste	Backtracking	Laufzeit	Grenze q
37	4	4,18s	1

A.5.4 Pelletiers 61. Problem

Zu zeigen:

$$\frac{(\forall x)(\forall y)(\forall z)(g(x,g(y,z)) \approx g(g(x,y),z))}{(\forall x)(\forall y)(\forall z)(\forall w)(g(x,g(y,g(z,w))) \approx g(g(g(x,y),z),w))}$$

Der Beweis zu diesem Beispiel ist prinzipiell sehr einfach zu führen. Zu zeigen ist der Abschluß des in Abbdildung A.6 dargestellten Tableaus.

Die Schwierigkeit liegt jedoch darin, daß das Assoziativgesetz zum Abschluß des Astes dreimal angewendet werden muß, und zwar mit verschiedenen Substitutionen für die in ihm enthaltenen freien Variablen. Wäre die Gleichung nicht als universell bezügliche der Variablen erkannt, könnte das Tableau nicht geschlossen werden. Dies wäre dann nur möglich, wenn drei Ausprägungen des Assoziativgesetzes zur Verfügung stünden, also für $q \geq 3$. Da die Universalität der Gleichung aber erkannt wird, genügt hier $q = 1$:

```
Beginning search for closure with equality of branch b1
```

$$\mathsf{T}\,(\forall x)(\forall y)(\forall z)(g(x,g(y,z)) \approx g(g(x,y),z))$$
$$|$$
$$\mathsf{T}\,(g(x_1,g(y_2,z_3)) \approx g(g(x_1,y_2),z_3))$$
$$|$$
$$\mathsf{F}\,(\forall x)(\forall y)(\forall z)(\forall w)(g(x,g(y,g(z,w))) \approx g(g(g(x,y),z),w))$$
$$|$$
$$\mathsf{F}\,(\forall y)(\forall z)(\forall w)(g(f_1,g(y,g(z,w))) \approx g(g(g(f_1,y),z),w))$$
$$|$$
$$\mathsf{F}\,(\forall z)(\forall w)(g(f_1,g(f_2,g(z,w))) \approx g(g(g(f_1,f_2),z),w))$$
$$|$$
$$\mathsf{F}\,(\forall w)(g(f_1,g(f_2,g(f_3,w))) \approx g(g(g(f_1,f_2),f_3),w))$$
$$|$$
$$\mathsf{F}\,(g(f_1,g(f_2,g(f_3,f_4))) \approx g(g(g(f_1,f_2),f_3),f_4))$$

Abbildung A.6: Tableau für Pelletiers 61. Problem.

```
Equalities extracted from Branch:
1:  [_1,_2,_3]  g(_1,g(_2,_3))  =  g(g(_1,_2),_3)
2:  [_1,_2,_3]  g(g(_1,_2),_3)  =  g(_1,g(_2,_3))

Disjunctions extracted from Branch:
(1,1): g(f1,g(f2,g(f3,f4)))  \=  g(g(g(f1,f2),f3),f4)  (expdbl.)

(1,1,l,0): (0,7) g(f1,g(f2,g(f3,f4))) []
(1,1,r,0): (0,7) g(g(g(f1,f2),f3),f4) []


(1,1,l,0): (0,7) g(f1,g(f2,g(f3,f4)))  -1->
                 (1,1,l,1): (0,7) g(g(f1,f2),g(f3,f4)) []
(1,1,l,0): (0,7) g(f1,g(f2,g(f3,f4)))  -1->
                 (1,1,l,2): (0,7) g(f1,g(g(f2,f3),f4)) []
(1,1,r,0): (0,7) g(g(g(f1,f2),f3),f4)  -2->
                 (1,1,r,1): (0,7) g(g(f1,f2),g(f3,f4)) []
(1,1,r,0): (0,7) g(g(g(f1,f2),f3),f4)  -2->
                 (1,1,r,2): (0,7) g(g(f1,g(f2,f3)),f4) []
Ineq. (1,1)  closed by  g(g(f1,f2),g(f3,f4))  =
                        g(g(f1,f2),g(f3,f4))  []
Disj. (1)  closed with  []
----------------------------------------------

-------------------- PROOF --------------------
```

Geschl. Äste	Backtracking	Laufzeit	Grenze q
1	0	0,33s	1

Wird die Universalität der Gleichung nicht beachtet, beträgt die Laufzeit 1,00s.

A.6 Beispiele aus der Gruppentheorie

Zum Abschluß zwei Beispiele aus der Gruppentheorie. Zu zeigen ist, daß aus den Axiomen

$$\begin{array}{l}(\forall x)(e \cdot x \approx x)\\(\forall x)(x \cdot x^{-1} \approx e)\\(\forall x)(x^{-1} \cdot x \approx e)\\(\forall x)(\forall y)(\forall z)(x \cdot (y \cdot z) \approx (x \cdot y) \cdot z)\end{array}$$

zunächst das von den anderen abhängige Axiom

$$(\forall x)(x \cdot e \approx x)$$

ableitbar ist, und dann mit dessen Hilfe das Theorem

$$(\forall x)(\forall y)(\forall z)(x \cdot y \approx z \cdot y \supset x \approx z),$$

also die Links–Eindeutigkeit der Verknüpfung in einer Gruppe.

Die Suche nach den Beweisen ist zu lang um sie hier wiederzugeben. In beiden Fällen besteht das Tableau aus nur einem Ast und alle Gleichungen sind universell bezüglich der Variablen, die sie enthalten. Das Problem besteht hier also „nur“ darin, die richtigen Ableitungsketten zu finden. Zum Beweis werden 88 bzw. 702 Terme abgeleitet.

Die Statistik zur Ableitung des zusätzlichen Axioms:

Geschl. Äste	Backtracking	Laufzeit	Grenze q
1	0	8,03s	1

Gefunden wird die Ableitungskette:

$$\begin{array}{rcl}x \cdot e & = & x \cdot (x^{-1} \cdot x)\\ & = & (x \cdot x^{-1}) \cdot x\\ & = & e \cdot x\\ & = & x\end{array}$$

Die Statistik zum Beweis der Eindeutigkeit der Verknüpfung:

Geschl. Äste	Backtracking	Laufzeit	Grenze q
1	0	389s	1

und die gefundene Ableitungskette:

$$\begin{array}{rcl}x & = & x \cdot e\\ & = & x \cdot (y \cdot y^{-1})\\ & = & (x \cdot y) \cdot y^{-1}\\ & = & (z \cdot y) \cdot y^{-1}\\ & = & z \cdot (y \cdot y^{-1})\\ & = & z \cdot e\\ & = & z\end{array}$$

Literaturverzeichnis

[Fitting, 1990] Melvin C. Fitting. *First–Order Logic and Automated Theorem Proving.* Springer, New York, 1990.

[Hähnle & Schmitt, 1991] Reiner Hähnle & Peter H. Schmitt. The liberalized δ–rule in free variable semantic tableaux. *To appear*, 1991.

[Hähnle, 1990a] Reiner Hähnle. Spezifikation eines Theorembeweisers für dreiwertige First–Order Logik. IWBS Report 136, Wissenschaftliches Zentrum, IWBS, IBM Deutschland, 1990.

[Hähnle, 1990b] Reiner Hähnle. Towards an efficient tableau proof procedure for multiple–valued logics. In *Proceedings Workshop on Computer Science Logic, Heidelberg.* Springer, LNCS 533, 1990.

[Hähnle, 1991] Reiner Hähnle. Uniform notation of tableau rules for multiple–valued logics. In *Proceedings International Symposium on Multiple–Valued Logic, Victoria.* IEEE Press, 1991.

[Jeffrey, 1967] R. C. Jeffrey. *Formal Logic: Its Scope and Limits.* McGraw Hill, 1967.

[Pelletier, 1986] Francis Jeffry Pelletier. Seventy–five problems for testing automatic theorem provers. *Journal of Automated Reasoning*, 2:191–216, 1986.

[Popplestone, 1967] R. J. Popplestone. Beth–tree methods in automatic theorem proving. In *Machine Intelligence*, volume 1, pages 31–46. Oliver and Boyd, 1967.

[Reeves, 1987] Steve V. Reeves. Adding equality to semantic tableau. *Journal of Automated Reasoning*, 3:225–246, 1987.