Technische Universität München
Institut für Informatik

Konfluenz bei Termersetzungssystemen auf λ-Termen

Diplomarbeit
von
Richard Mayr

Aufgabensteller und Betreuer: Prof. Dr. Tobias Nipkow, Ph.D.
Abgabedatum: 15.2.1994

Ich versichere, daß ich diese Diplomarbeit selbständig verfaßt und nur die angegebenen Quellen und Hilfsmittel verwendet habe.

München, den 7. Februar 1994, ...

Zusammenfassung

Konfluenz ist eine häufig gewünschte Eigenschaft von Termersetzungssystemen. Daher ist es nötig, Kriterien für die Konfluenz eines Termersetzungssystems zu finden. Viele der bisherigen Ergebnisse sind jedoch nur für Termersetzungssysteme auf Termen *erster* Ordnung bewiesen. Hier werden einige dieser Ergebnisse auf Termersetzungssysteme *höherer* Ordnung verallgemeinert. Es werden dabei Termersetzungssysteme auf einfach typisierten λ-Termen mit eingeschränkten linken Seiten betrachtet.

Inhaltsverzeichnis

Kapitel 1

Einleitung

In vielen Anwendungen wie funktionalen Programmiersprachen und Theorembeweisern werden Ergebnisse berechnet, indem Terme durch Anwendung von bestimmten Regeln reduziert werden. Eine Form solcher Regelsysteme sind die sogenannten Termersetzungssysteme. Da diese schon seit längerem studiert werden sind eine Vielzahl von Ergebnissen dafür verfügbar. Ein großer Teil dieser Ergebnisse ist jedoch nur für Termersetzungssysteme erster Ordnung bewiesen. Bei Termersetzungssystemen erster Ordnung sind alle Variablen von Basistypen wie z.B. $int, bool, \ldots$ nicht aber von Funktionstypen wie z.B. $(int \to int) \to bool$.

Ein Ansatz zur Verallgemeinerung sind die *kombinatorischen Reduktionssysteme* (kurz: CRS) die von Klop [11] eingeführt wurden. Ein neuerer Ansatz dazu sind die *Termersetzungssysteme höherer Ordnung* (kurz: HRS, für *higher-order rewrite systems*), von Nipkow [2] die den einfach typisierten λ-Kalkül als Metasprache zur Beschreibung der Reduktion von Termen mit gebundenen Variablen verwenden. Abgesehen von Kapitel 7 werden in dieser Arbeit nur HRS betrachtet.

Termersetzungssysteme höherer Ordnung haben die gleiche logische Basis wie das Theorembeweisersystem Isabelle [15] und die Programmiersprache λProlog [16, 17], die ebenfalls Terme mit gebundenen Variablen manipulieren.

Termersetzungssysteme können als Berechnungsmechanismus verstanden werden, und tatsächlich ist jede berechenbare Funktion durch ein Termersetzungssystem darstellbar. Insbesondere ist die Terminierung von Termersetzungssystemen im allgemeinsten Fall nicht entscheidbar, da sonst das Halteproblem entscheidbar wäre. Außerdem ist die Berechnung durch ein Termersetzungssystem nichtdeterministisch, da auf einen Term häufig mehrere Termersetzungsregeln anwendbar sind. Die Eigenschaft der Konfluenz ist für Termersetzungssysteme deswegen wichtig, weil aus ihr die Eindeutigkeit von Normalformen folgt. Das heißt, ein konfluentes Termersetzungssystem kann auf einen Eingabeterm angewandt bei zwei verschiedenen terminierenden Berechnungspfaden nicht zwei verschiedene Ergebnisse liefern. Anders ausgedrückt kann man sagen, daß konfluente Termersetzungssysteme (partielle) Funktionen berechnen (auch wenn sie nicht für alle

Eingabeterme terminieren).

Im Rest dieses Kapitels werden die verwendete Metasprache und einige notationelle Konventionen definiert. Im zweiten Kapitel werden einige aus Arbeiten von Nipkow [2, 3] stammende Ergebnisse aufgeführt, da sie in den folgenden Kapiteln verwendet werden. Im dritten Kapitel werden diese Ergebnisse verwendet, um einen Satz über lokale Konfluenz von HRS zu beweisen. Im vierten Kapitel werden einige Ergebnisse von Huet [1] über die Konfluenz modulo Gleichungssystemen von Termersetzungssystemen erster Ordnung auf Termersetzungssysteme höherer Ordnung verallgemeinert. Im fünften Kapitel werden die Beziehungen zwischen Termersetzung höherer Ordnung und Gleichheit modulo eines Termersetzungssystems untersucht. Das sechste Kapitel befaßt sich mit der Konfluenz von nicht-noetherschen Termersetzungssystemen auf λ-Termen. Hier wird ein Satz von Huet [1] mit Einschränkungen auf Termersetzungssysteme höherer Ordnung verallgemeinert, sowie ein Ergebnis von Takahashi [6] in etwas veränderter Form dargestellt. Im siebten Kapitel werden die kombinatorischen Reduktionssysteme (CRS) kurz vorgestellt und mit den Termersetzungssystemen höherer Ordnung (HRS) verglichen. Im letzten Kapitel werden die Ergebnisse zusammengefaßt und einige noch offene Probleme dargestellt.

1.1 Grundlagen und notationelle Konventionen

Es wird die Standardtheorie des einfach typisierten λ-Kalküls verwendet (siehe [5]). Die Notation ist im wesentlichen konsistent mit der von Nipkow [3]. Im folgenden wird noch einmal das Wichtigste aufgeführt.

1.1.1 Typen und Terme

Sei $\mathcal{T}$ die Menge aller Typen und $\tau \in \mathcal{T}$. $\mathcal{T}$ wird durch den Funktionskonstruktor $\rightarrow$ aus den **Basistypen** (z.B. *nat, bool*) erzeugt. Der Funktionskonstruktor ist rechtsassoziativ, d.h. $\tau_1 \rightarrow \tau_2 \rightarrow \tau_3 = \tau_1 \rightarrow (\tau_2 \rightarrow \tau_3)$. Falls τ ein Basistyp ist wird statt $\tau_1 \rightarrow \ldots \rightarrow \tau_n \rightarrow \tau$ auch $\overline{\tau_n} \rightarrow \tau$ geschrieben. Falls τ ein Basistyp ist wird die **Ordnung** des Typs $\varphi = \tau_1 \rightarrow \ldots \rightarrow \tau_n \rightarrow \tau$ definiert als

$$Ord(\varphi) = \begin{cases} 1 & \text{falls } n = 0 \text{, d.h. } \varphi \text{ ist ein Basistyp} \\ 1 + max(Ord(\tau_1), \ldots, Ord(\tau_n)) & \text{sonst} \end{cases}$$

Terme werden aus einer Menge von typisierten **Variablen** $V = \bigcup_{\tau \in \mathcal{T}} V_\tau$ und einer Menge von typisierten **Konstanten** $C = \bigcup_{\tau \in \mathcal{T}} C_\tau$ aufgebaut, wobei $V_\tau \cap V_{\tau'} = C_\tau \cap C_{\tau'} = \emptyset$ für $\tau \neq \tau'$. Gebundene Variablen werden als Kleinbuchstaben x, y und z, freie Variable als Großbuchstaben F, G, H, ... und Konstanten als c geschrieben. **Atome** sind Konstanten oder Variable und werden als a geschrieben. Terme werden als l, r, s, t, u und v geschrieben. Ein λ-Term ist entweder ein Atom a, eine Abstraktion $\lambda x.t$ oder eine Applikation $(s\ t)$, wobei s und t wieder λ-Terme sind.

Im folgenden seien alle λ-Terme einfach typisiert. Wenn ein Term t den Typ τ hat, so wird dies als $t : \tau$ geschrieben. Die induktive Definition der einfach typisierten λ-Terme ist wie folgt.

$$\frac{x \in V_\tau}{x : \tau} \quad \frac{c \in C_\tau}{c : \tau} \quad \frac{s : \tau \rightarrow \tau' \quad t : \tau}{(s\ t) : \tau'} \quad \frac{x : \tau \quad s : \tau'}{(\lambda x.s) : \tau \rightarrow \tau'}$$

Wenn ein λ-Term t den Typ τ hat und τ von der Ordnung n ist, so wird auch gesagt, daß t die Ordnung n hat. Eine **Sprache der Ordnung n** darf nur

- Funktionskonstanten der Ordnung $\leq n + 1$ und
- Variablen der Ordnung $\leq n$

enthalten.
Im typisierten λ-Kalkül lassen sich Terme beliebiger Ordnung konstruieren. Daher ist eine Sprache n-ter Ordnung (und insbesondere erster Ordnung) eine echte Untermenge des typisierten λ-Kalküls.

Funktionen mit mehreren Argumenten werden durch das sogenannte **"currying"** dargestellt. Dabei liefert eine Funktion mit n Argumenten nach Anwendung auf das erste Argument als Ergebnis eine Funktion die $n - 1$ Argumente nimmt. Eine Funktion f, die drei Werte vom Typ *int* als Argumente nimmt und einen Wert vom Typ *boolean* als Ergebnis liefert, hat dann den Typ $f : int \to int \to int \to bool$. (Beachte, daß der Funktionskonstruktor $\to$ rechtsassoziativ ist.) Statt der normalen Schreibweise der Funktionsanwendung $(((f\ c_1)\ c_2)\ c_3)$ wird vereinfachend auch $f(c_1, c_2, c_3)$ geschrieben.

Bei Funktionstermen wird statt $\lambda x_1 \ldots \lambda x_n.s$ auch $\lambda x_1, \ldots, x_n.s$ oder $\lambda\overline{x_n}.s$ geschrieben, wobei die x_i alle verschieden sein sollen. Ebenso wird anstatt von $(\ldots(t\ u_1)\ldots)u_n$ auch $t(u_1, \ldots, u_n)$ oder nur $t(\overline{u_n})$ geschrieben. Die **freien Variablen** eines Terms s seien $fv(s)$, die **gebundenen Variablen** von s seien $bv(s)$. Sie sind definiert durch

$$\begin{array}{rcll} fv(x) & = & x & \text{für } x \in V \\ fv(c) & = & \emptyset & \text{für } c \in C \\ fv((s\ t)) & = & fv(s) \cup fv(t) & \\ fv(\lambda x.t) & = & fv(t) - \{x\} & \end{array}$$

und

$$\begin{array}{rcll} bv(x) & = & \emptyset & \text{für } x \in V \\ bv(c) & = & \emptyset & \text{für } c \in C \\ bv((s\ t)) & = & bv(s) \cup bv(t) & \\ bv(\lambda x.t) & = & bv(t) \cup \{x\} & \end{array}$$

1.1.2 Die α-, β- und η-Konversion

Zwei λ-Terme s und t sind α-äquivalent, wenn sie durch Umbenennung ihrer gebundenen Variablen ineinander überführt werden können. Man schreibt dafür $s =_\alpha t$. Der Übergang von s nach t bzw. t nach s heißt dann α-Konversion.

Durch die β-Reduktion $((\lambda x.s)\ t) \to_\beta s[x := t]$ wird ein Term der Form $((\lambda x.s)\ t)$ in den Term $s[x := t]$ überführt, wobei $s[x := t]$ der Term ist, der aus s entsteht, wenn in s für die Variable x der Term t eingesetzt wird. Die β-Reduktion bedeutet also das Einsetzen der Argumente in die Parameter einer Funktion. Für den untypisierten λ-Kalkül ist die β-Reduktion nicht terminierend, wie schon das folgende einfache Beispiel zeigt.

$$((\lambda x.(x\ x))\ (\lambda x.(x\ x))) \to_\beta ((\lambda x.(x\ x))\ (\lambda x.(x\ x))) \to_\beta \ldots$$

Im Gegensatz dazu ist die β-Reduktion für den einfach typisierten λ-Kalkül konfluent und terminierend und somit stark normalisierend. Da hier nur der einfach typisierte λ-Kalkül betrachtet wird ist für jeden λ-Term t seine β-Normalform definiert. Die **β-Normalform** eines λ-Terms t wird als $t{\downarrow_\beta}$ geschrieben. $t{\downarrow_\beta}$ ist dann von der Form $\lambda\overline{x_n}.a(\overline{u_m})$, wobei das Atom a der **Kopf** von t genannt wird und alle u_i in β-Normalform sind. Wenn $s\ (\leftrightarrow_\beta)^*\ t$ gilt, dann heißen s und t β-gleich. Dies wird auch als $s =_\beta t$ geschrieben. Die Relation $=_\beta$ heißt auch β-Konversion. Im einfach typisierten λ-Kalkül gilt $s =_\beta t \Leftrightarrow s{\downarrow_\beta} = t{\downarrow_\beta}$.

Eine weitere Umformung ist die η-Reduktion, die als $\lambda x.F(x) \to_\eta F$ definiert ist. Die Umkehrung $\lambda x.F(x) \leftarrow_\eta F$ heißt η-Expansion. Wenn $s\ (\leftrightarrow_\eta)^*\ t$ gilt, dann heißen s und t η-gleich. Dies wird als $s =_\eta t$ geschrieben. Die Relation $=_\eta$ heißt auch η-Konversion.

Für einen Term t in β-Normalform, d.h. $t = t{\downarrow_\beta}$, ist die **η-expandierte Form** von t definiert als

$$t{\uparrow^\eta} = \lambda\overline{x_{n+k}}.a(\overline{u_m}{\uparrow^\eta}, x_{n+1}{\uparrow^\eta}, \ldots, x_{n+k}{\uparrow^\eta})$$

wobei $t : \overline{\tau_{n+k}} \to \tau$ und $x_{n+1}, \ldots, x_{n+k} \notin fv(\overline{u_m})$. Statt $t{\downarrow_\beta}{\uparrow_\eta}$ wird auch $t\updownarrow^\eta_\beta$ geschrieben. Ein λ-Term t ist in **langer $\beta\eta$-Normalform**, falls $t = t\updownarrow^\eta_\beta$.

Die $\beta\eta$-Gleichheit ist definiert durch

$$s =_{\beta\eta} t \quad :\Leftrightarrow \quad s\ (\leftrightarrow_\beta \cup \leftrightarrow_\eta)^*\ t$$

Im einfach typisierten λ-Kalkül gilt $s =_{\beta\eta} t \ \Leftrightarrow\ s\updownarrow^\eta_\beta = t\updownarrow^\eta_\beta$.

1.1.3 Subterme und Positionen

Terme lassen sich auch als Bäume darstellen, wobei man Subterme mit Positionen im Baum identifizieren kann. Diese **Positionen** im Baum werden als Pfade von der Wurzel bis zum Subterm aufgefaßt. Näheres dazu steht in [1, 7]. Hier wird nur kurz die Notation aufgeführt.

Sei $Pos(t)$ die Menge der Positionen in einem Term t. Diese werden als Sequenzen natürlicher Zahlen dargestellt, d.h. $Pos(t) \subseteq \mathbb{N}^*$. Die Anfangsposition wird als die leere Sequenz ϵ dargestellt. Sei $p \in Pos(t)$, dann ist t/p der Subterm von t an der Position p. Positionen können aneinandergefügt werden. Für $q \in Pos(t/p)$ ist $p{\cdot}q \in Pos(t)$, wobei $t/p{\cdot}q = (t/p)/q$. Dies wird in der Abbildung 1.1 dargestellt. Für eine Menge von Positionen W und eine Position p ist $p \cdot W := \{p \cdot u \mid u \in W\}$.

Der Term der aus t entsteht indem man an der Position p den Subterm t/p durch den Term u ersetzt wird als $t[u]_p$ geschrieben. Wenn an mehreren Stellen gleichzeitig Subterme ersetzt werden sollen, so wird dies als $t[p \leftarrow u \mid B]$ geschrieben, wobei p eine Position in t, u der Term der t/p ersetzen soll und B eine Bedingung ist die erfüllt sein muß, damit die Ersetzung stattfindet. Ein Beispiel dafür ist $t[p_i \leftarrow u_i \mid i = 1, \ldots, n]$.

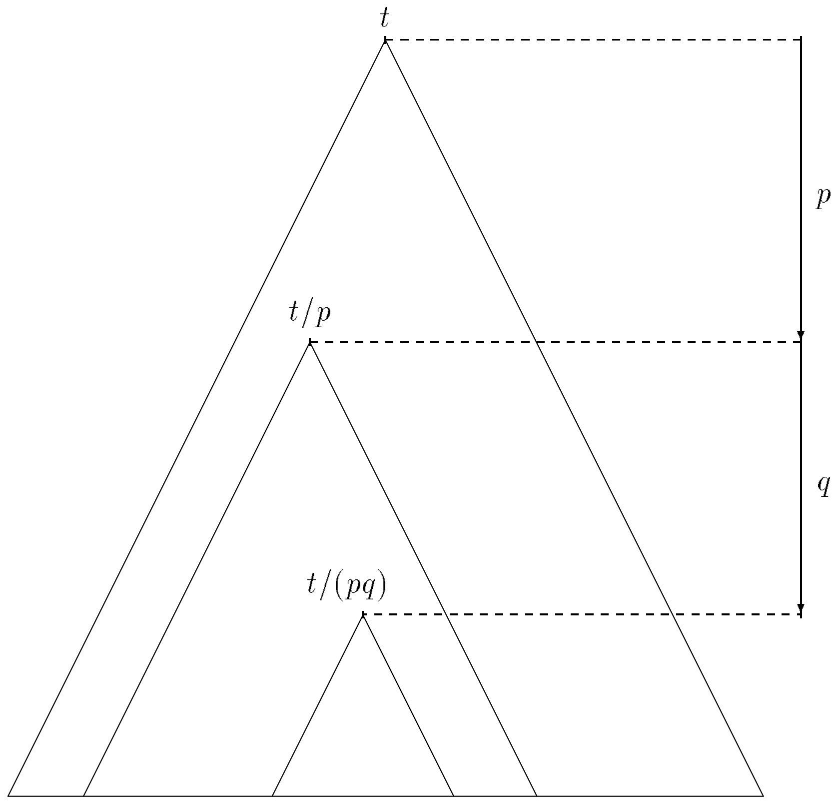
Abstraktion ($\lambda x.s$) und Applikation ($(s\ t)$) erzeugen folgende Bäume:

$$\begin{array}{c} \lambda x \\ | \\ s \end{array} \qquad\qquad \begin{array}{c} . \\ /\ \backslash \\ s \quad t \end{array}$$

Formal wird dies definiert durch

$$\begin{array}{rcl} t/\epsilon & = & t \\ (s\ t)/1p & = & s/p \\ (s\ t)/2p & = & t/p \\ (\lambda x.s)/1p & = & s/p \end{array}$$

Abbildung 1.1: Terme und Subterme



Positionen in λ-Termen sind also Sequenzen über $\{1,2\}$. Außerdem ist zu bemerken, daß eine in t gebundene Variable kein eigener Subterm von t ist und daher nicht mit der Notation t/p erreicht werden kann.

Die Menge $Pos(t)$ der Positionen in einem λ-Term t ist partiell geordnet durch die Präfixordnung $\leq$. Für zwei Positionen $p_1, p_2 \in Pos(t)$ gilt $p_1 \leq p_2$ genau dann, wenn p_1 ein Präfix von p_2 ist. Wenn zwei Positionen p_1 und p_2 unvergleichbar sind, d.h. wenn weder p_1 ein Präfix von p_2, noch p_2 ein Präfix von p_1 ist, so schreibt man $p_1 \parallel p_2$, um dies anzuzeigen.

Für zwei Positionen $p_1, p_2 \in Pos(t)$ mit $p_1 \leq p_2$ ist die Position $p_2/p_1 \in Pos(t/p_1)$ definiert als die Sequenz, die aus p_2 entsteht, wenn das Präfix p_1 entfernt wird.

Zur Vereinfachung wird im weiteren Barendregts Konvention zur Behandlung der λ-Terme angenommen: α-äquivalente Terme werden identifiziert und es wird angenommen, daß die in einem Zusammenhang vorkommenden Mengen der freien- und der gebundenen Variablen disjunkt sind. Ausnahmen von dieser Konvention treten dann auf, wenn auf Subterme zugegriffen wird, da eine im Term t gebundene Variable in einem Subterm t/p auch frei sein kann.

1.1.4 Länge von Termen

Da in den Beweisen mehrerer Lemmata eine Induktion über die Größe eines λ-Terms geführt wird ist es nötig, ein Längenmaß für λ-Terme zu definieren.

Definition 1.1.1 (Länge eines Terms) Das Längenmaß $L : \mathcal{T} \mapsto \mathbb{N}$ wird induktiv über den Termaufbau definiert durch:

$$\begin{aligned} L(a) &:= 1 \\ L((s\ t)) &:= L(s) + L(t) \\ L(\lambda x.t) &:= 1 + L(t) \end{aligned}$$

wobei a ein Atom ist.

Bemerkungen:

- Aus der Definition folgt direkt, daß $L(t) > 0$ gilt für jeden λ-Term t.
- Ebenso gilt für jeden λ-Term t und $\{p_1, \ldots, p_n\} \subseteq Pos(t)$ mit $p_i \parallel p_j$ für $i \neq j$ die Ungleichung $\sum_{i=1}^{n} L(t/p_i) \leq L(t)$. Die Abbildung 1.2 illustriert diesen Sachverhalt.
- Insbesondere gilt für jeden λ-Term t und jede Position $p \in Pos(t)$ die Ungleichung $L(t) \geq L(t/p)$.

1.1.5 Substitutionen

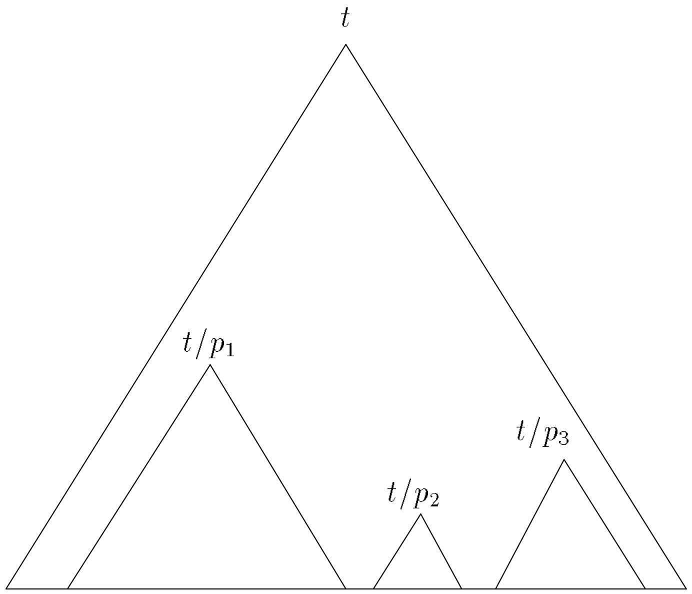
Substitutionen sind *endliche* Abbildungen von der Menge der Variablen V in die Menge der Terme $\mathcal{T}$, d.h. für jede Substitution θ gilt $\theta(x) \neq x$ für nur endlich viele Variablen $x \in V$. Dabei müssen der Typ der Variable x und des Terms $\theta(x)$ übereinstimmen. Substitutionen werden meist mit den Buchstaben θ, σ und δ bezeichnet. Für $\sigma = \{x_1 \mapsto t_1, \ldots, x_n \mapsto t_n\}$ ist der Definitionsbereich $Dom(\theta) = \{x_1, \ldots, x_n\}$ und der Wertebereich $Cod(\theta) = \{t_1, \ldots, t_n\}$.

Konvention: Für jede Substitution θ seien alle Elemente von $Cod(\theta)$ in langer $\beta\eta$-Normalform.

Eine **Umbenennung** ρ ist eine injektive Substitution mit $Cod(\rho) \subset V$ und $Dom(\rho) \cap Cod(\rho) = \emptyset$. Umbenennungen werden immer mit ρ bezeichnet.

Seien θ_1 und θ_2 Substitutionen. Dann sei $\theta_1 + \theta_2$ eine Substitution mit $Dom(\theta_1 + \theta_2) = Dom(\theta_1) \cup \mathcal{D}om(\theta_2)$ die definiert ist durch
$$(\theta_1 + \theta_2)(F) = \begin{cases} \theta_2(F) & \text{falls } F \in Dom(\theta_2) \\ \theta_1(F) & \text{sonst} \end{cases}$$

Abbildung 1.2: Parallele Subterme



Wegen der oben erwähnten Konvention für die Variablen kann die Anwendung jeder Substitution einfach auf Terme fortgesetzt werden. Sei eine Substitution $\theta : V \mapsto \mathcal{T}$ gegeben. Dann ist ihre Erweiterung $\overline{\theta} : \mathcal{T} \mapsto \mathcal{T}$ definiert als:

$$\begin{array}{rcl} \overline{\theta}(c) & = & c \\ \overline{\theta}(F) & = & \theta(F) \\ \overline{\theta}(\ (s\ t)\ ) & = & (\overline{\theta}(s)\ \overline{\theta}(t)) \\ \overline{\theta}(\lambda x.s) & = & \lambda x.\overline{\theta}(s) \end{array}$$

Konvention: Im folgenden ist, wenn von einer Substitution gesprochen wird, stets ihre Erweiterung gemeint.

Zwei Terme s und t heißen **unifizierbar**, wenn es eine Substitution θ gibt, so daß $\theta(s)$ und $\theta(t)$ $\beta\eta$-äquivalent sind. Da die β-Reduktion im typisierten λ-Kalkül stark normalisierend ist, gilt dies genau dann, wenn $\theta(s)\updownarrow_\beta^\eta = \theta(t)\updownarrow_\beta^\eta$.

1.1.6 Einsetzungen

Sei s ein λ-Term in dem die Variable F frei vorkommt, d.h. $F \in fv(s)$ und t ein weiterer λ-Term. Dann bezeichnet $s[F := t]$ den λ-Term der aus s entsteht, wenn in s der Term t für die Variable F eingesetzt wird. Im Gegensatz zu der für die Substitutionen gewählten Konvention wird hier *nicht* vorausgesetzt, daß der Term t in langer $\beta\eta$-Normalform ist. Davon abgesehen ist $s[F := t]$ definiert als $\theta(s)$ mit $\theta := \{F \mapsto t\}$.

1.2 Termersetzungssysteme auf λ-Termen

Termersetzungssysteme auf λ-Termen (auch *HRS=Higher-Order Rewrite Systems* genannt) sind eine Verallgemeinerung von Termersetzungssystemen auf Termen erster Ordnung (auch *TRS=Term Rewriting Systems* genannt). Ausführlicheres dazu steht in [2, 3]. Hier werden nur die wichtigsten Definitionen und Lemmata aufgeführt.

Da das Problem der Unifizierbarkeit zweier λ-Terme im allgemeinen Fall unentscheidbar ist [8] wird eine Einschränkung von λ-Termen definiert, mit der sich besser arbeiten läßt.

Definition 1.2.1 (Pattern) Ein λ-Term t in β-Normalform ist ein **"Pattern"**, falls alle freien Variablen $F \in fv(t)$ nur in Subtermen von t von der Form $F(\overline{u_n})$ auftreten, wobei $\overline{u_n}$ η-äquivalent zu einer Liste von *verschiedenen* gebundenen Variablen ist.

Beispiel 1.2.2 Patterns und nicht-Patterns.

- Beispiele für Patterns sind $\lambda x.c(x)$, X, $\lambda x.F(\lambda z.x(z))$ und $\lambda x, y.F(y, x)$.
- Beispiele für Terme, die keine Patterns sind: $F(c)$, $\lambda x.F(x, x)$, $\lambda x, y.F(y, c)$ und $\lambda x.G(H(x))$.

Das folgende wichtige Ergebnis über die Unifikation von Patterns wurde von Dale Miller [4] bewiesen.

Satz 1.2.3 *Es ist entscheidbar, ob zwei Patterns unifizierbar sind. Wenn sie unifizierbar sind kann ein allgemeinster Unifikator berechnet werden.*

Eine vereinfachte Version von Millers Unifikationsalgorithmus wurde von Nipkow in [2] beschrieben. Er wurde von ihm in [9] im Hinblick auf eine praktische Implementierung weiterentwickelt.

Definition 1.2.4 Eine **Termersetzungsregel** ist ein geordnetes Paar von λ-Termen l und r das als $l \to r$ geschrieben wird und für das gilt:

- l ist ein Pattern aber nicht η-äquivalent zu einer freien Variable.
- l und r sind in langer $\beta\eta$-Normalform und vom gleichen Typ, der ein Basistyp sein muß.

- $fv(l) \supseteq fv(r)$

Ein **Termersetzungssystem höherer Ordnung** (kurz: **HRS**) ist eine endliche Menge von Termersetzungsregeln und wird meist mit dem Buchstaben R bezeichnet. Ein HRS R induziert eine Relation $\rightarrow_R$ auf λ-Termen in langer $\beta\eta$-Normalform:

$$s \rightarrow_R t \ :\Leftrightarrow \ \exists (l \rightarrow r) \in R, p \in Pos(s), \theta. \ s/p = \theta(l)\downarrow_\beta \ \wedge \ t = s[\theta(r)\downarrow_\beta]_p$$

Der Subterm s/p von s an der Position p wird **Redex** (kurz für: reducible expression) genannt. Der Term t/p wird das **Kontraktum** des Redex s/p genannt. Der Übergang von s nach t wird **Reduktionsschritt** genannt.

Einige Erläuterungen zu dieser Definition:

- Die Relation $\rightarrow_R$ ist berechenbar, da nach Satz 1.2.3 die Substitution θ berechenbar ist.

- Es ist keine Einschränkung, daß die Relation $\rightarrow_R$ nur auf Termen in langer $\beta\eta$-Normalform definiert ist, da jeder Term eine eindeutige lange $\beta\eta$-Normalform hat.

- In obiger Definition von $\rightarrow_R$ ist der Term $t = s[\theta(r)\downarrow_\beta]_p$ in langer $\beta\eta$-Normalform unter der Voraussetzung, daß r, s und die Elemente von $Cod(\theta)$ es sind: t ist in β-Normalform, weil $\theta(r)\downarrow_\beta$ von Basistyp ist; t ist in η-expandierter Form, weil $\theta(r)\downarrow_\beta$ es ist.

- Es ist möglich die Relation $\rightarrow_R$ in gleicher Weise auch für HRS zu definieren, bei denen nicht alle Regeln $l \rightarrow r$ von Basistyp sind. Die meisten der in den folgenden Kapiteln bewiesenen Ergebnisse gelten jedoch für HRS ohne diese Einschränkung nicht mehr (siehe dazu [3]).

- Wenn aus dem aktuellen Kontext klar ist welches HRS R gemeint ist wird statt $\rightarrow_R$ nur $\rightarrow$ geschrieben.

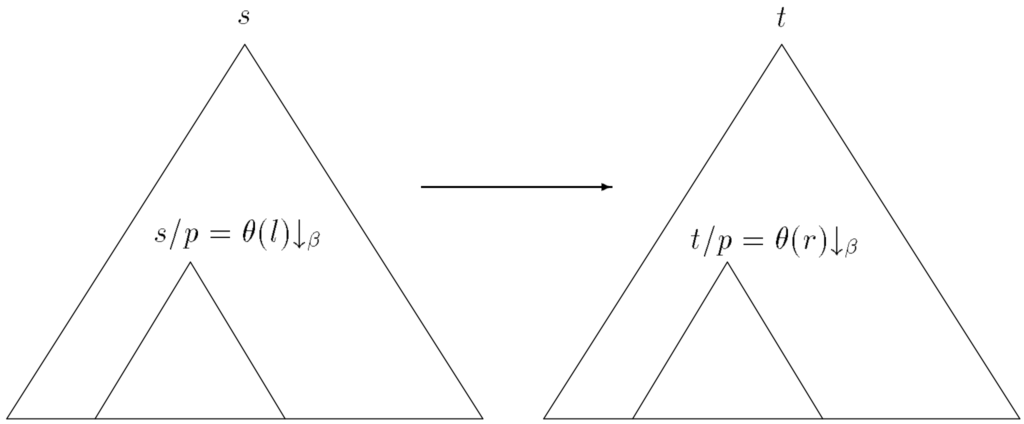
Die Abbildung 1.3 illustriert den Übergang von s nach t:

Definition 1.2.5 (Linearität) Als Spezialfälle von HRS werden definiert:

- Eine Termersetzungsregel $(l \rightarrow r)$ heißt **linkslinear**, wenn jede freie Variable $F \in fv(l)$ nur *einmal* in l vorkommt. Ein HRS R heißt linkslinear, wenn alle Regeln in R linkslinear sind.

- Eine Termersetzungsregel $(l \rightarrow r)$ heißt **rechtslinear**, wenn jede freie Variable $F \in fv(r)$ nur *einmal* in r vorkommt. Ein HRS R heißt rechtslinear, wenn alle Regeln in R rechtslinear sind.

Abbildung 1.3: Reduktion von Termen



Beispiel 1.2.6 Zur Darstellung der Syntax des reinen λ-Kalküls werden neben den Mengen der Variablen und Konstanten nur ein Typ *term* der λ-Terme und zwei Konstanten für die Abstraktion und die Applikation benötigt.

$$\begin{array}{rcl} abs & : & (term \rightarrow term) \rightarrow term \\ app & : & term \rightarrow term \rightarrow term \end{array}$$

Damit lassen sich die β- und η-Konversion durch Termersetzungsregeln darstellen, die sowohl links- als auch rechtslinear sind:

$$\begin{array}{rcrcl} \beta & : & app(abs(\lambda x.F(x)), S) & \rightarrow & F(S) \\ \eta & : & abs(\lambda x.app(H, x)) & \rightarrow & H \end{array}$$

Kapitel 2

Vorbereitung

In diesem Kapitel werden zuerst einige Definitionen und Lemmata aus der Relationentheorie aufgeführt, da sie in den folgenden Kapiteln verwendet werden. Danach werden einige wichtige Definitionen und Beweistechniken für HRS dargestellt. Diese werden dann im nächsten Kapitel im Beweis eines Ergebnisses von Nipkow [2] über lokale Konfluenz von HRS verwendet werden.

2.1 Definitionen und Lemmata aus der Relationentheorie

Da die folgenden Kapitel davon handeln, unter welchen Bedingungen die durch ein HRS R induzierte Relation $\rightarrow_R$ auf den λ-Termen die relationentheoretische Eigenschaft der Konfluenz bzw. Konfluenz modulo eines Gleichungssystems hat, ist es nötig diese Eigenschaften zu definieren und einige relationentheoretische Lemmata anzugeben, die in den Beweisen Verwendung finden. Die Beweise der folgenden Lemmata stehen in [1] bzw. in Büchern über Relationentheorie.

Definition 2.1.1 Sei $\rightarrow$ eine binäre Relation auf einer Menge M. Dann ist

ι die Identität auf M	$\iota = \{< x, x > \mid x \in M\}$
$\cdot$ die Relationskomposition	$\rightarrow_a \cdot \rightarrow_b = \{< x, z > \mid \exists y. x \rightarrow_a y \wedge y \rightarrow_b z\}$
$\rightarrow^{-1}$ die inverse Relation	$x \rightarrow^{-1} y :\Leftrightarrow y \rightarrow x$

Für jede Relation $\rightarrow$ auf M sei:

$\rightarrow^0$	$=$	ι	
$\rightarrow^i$	$=$	$\rightarrow \cdot \rightarrow^{i-1}\ \ \forall i > 0$	die i-fache Komposition von $\rightarrow$
$\rightarrow^=$	$=$	$\rightarrow \cup \iota$	die reflexive Hülle von $\rightarrow$
$\rightarrow^+$	$=$	$\cup_{i>0} \rightarrow^i$	die transitive Hülle von $\rightarrow$
$\rightarrow^*$	$=$	$\rightarrow^+ \cup \iota$	die reflexiv-transitive Hülle von $\rightarrow$
$\leftrightarrow$	$=$	$\rightarrow \cup \rightarrow^{-1}$	die symmetrische Hülle von $\rightarrow$

Sei $x \in M$ mit $\nexists y \in M.\ x \to y$, dann heißt x **$\to$-Normalform**. Falls ein $z \in M$ existiert mit $z \to^* x$, dann heißt x **$\to$-Normalform von** z. Für eine Relation $\to$ wird definiert

$x \downarrow y$	$:\Leftrightarrow$	$\exists z.\ x \to^* z \ \wedge\ y \to^* z$	x und y haben gemeinsamen Nachfolger
$x \uparrow y$	$:\Leftrightarrow$	$\exists z.\ z \to^* x \ \wedge\ z \to^* y$	x und y haben gemeinsamen Vorgänger

Definition 2.1.2 (Terminierend) Eine Relation $\to$ heißt terminierend oder noethersch, falls es keine unendlichen Folgen $x_1 \to x_2 \to \ldots$ gibt.

Definition 2.1.3 (Konfluenz) Eine Relation $\to$ heißt konfluent falls gilt:

$$\forall x, y.\ x \uparrow y \ \Rightarrow\ x \downarrow y.$$

Lemma 2.1.4 *Sei die Relation $\to$ auf der Menge M konfluent. Dann hat jedes Element aus M höchstens eine $\to$-Normalform.*

Definition 2.1.5 (Lokale Konfluenz) Eine Relation $\to$ ist lokal konfluent falls gilt:

$$\forall x, y, z.\ x \to y \ \wedge\ x \to z \ \Rightarrow\ y \downarrow z.$$

Lemma 2.1.6 *Eine terminierende Relation ist genau dann konfluent, wenn sie lokal konfluent ist.*

Definition 2.1.7 (Diamanteigenschaft) Eine Relation $\to$ hat die Diamanteigenschaft (diamond property) falls gilt:

$$x \to y \ \wedge\ x \to z \ \Rightarrow\ \exists u.\ y \to u \ \wedge\ z \to u$$

Lemma 2.1.8 *Seien $\to$ und $>$ Relationen auf der gleichen Menge, wobei $\to\ \subseteq\ >\ \subseteq\ \to^*$ gilt und $>$ die Diamanteigenschaft hat. Dann ist $\to$ konfluent.*

Definition 2.1.9 (Konfluenz modulo einer Äquivalenzrelation) Eine Relation $\to$ heißt konfluent modulo einer Äquivalenzrelation $\sim$ wenn gilt:

$$\forall x, y, x', y'.\ x \sim y \ \wedge\ x \xrightarrow{*} x' \ \wedge\ y \xrightarrow{*} y' \ \Rightarrow\ \exists x'', y''.\ x' \xrightarrow{*} x'' \ \wedge\ y' \xrightarrow{*} y'' \ \wedge\ x'' \sim y''$$

Lemma 2.1.10 *Sei $\to$ eine terminierende Relation auf einer Menge M und $N \subseteq M$ die Menge aller $\to$-Normalformen. Die Relation $\to$ ist genau dann konfluent modulo einer Äquivalenzrelation $\sim$ auf M wenn gilt:*

$$\forall x, y \in M\ \forall u, v \in N.\ x \equiv y \ \wedge\ x \xrightarrow{*} u \ \wedge\ y \xrightarrow{*} v \ \Rightarrow\ u \sim v$$

wobei $\equiv$ gleich $(\leftrightarrow \cup \sim)^*$ ist.

Definition 2.1.11 Sei $\rightarrow$ eine Relation auf einer Menge M und $\sim$ eine Äquivalenzrelation auf M. Die Bedingung, daß zwei Elemente $y, z \in M$ ein gemeinsames Redukt modulo $\sim$ haben wird definiert als:

$$y \overset{\sim}{\downarrow} z \;:\Leftrightarrow\; \exists u, v.\ y \overset{*}{\rightarrow} u \;\wedge\; z \overset{*}{\rightarrow} v \;\wedge\; u \sim v$$

Die speziellen Eigenschaften $\mathcal{A}$ und $\mathcal{B}$ sind definiert als:

$$\begin{aligned}\mathcal{A}: &\quad \forall x, y, z.\ x \rightarrow y \;\wedge\; x \rightarrow z \;\Rightarrow\; y \overset{\sim}{\downarrow} z\\ \mathcal{B}: &\quad \forall x, y, z.\ x \sim y \;\wedge\; x \rightarrow z \;\Rightarrow\; y \overset{\sim}{\downarrow} z\end{aligned}$$

Die Relation $\rightarrow$ heißt **lokal konfluent modulo** $\sim$ genau dann, wenn sie die Eigenschaften $\mathcal{A}$ und $\mathcal{B}$ erfüllt. Wenn die Relation $\sim$ von einer symmetrischen Relation $\vdash\!\dashv$ erzeugt wird, d.h. $\sim = \overset{*}{\vdash\!\dashv}$, sei die Eigenschaft $\mathcal{C}$ als Abschwächung von $\mathcal{B}$ definiert als

$$\mathcal{C}: \quad \forall x, y, z.\ x \vdash\!\dashv y \;\wedge\; x \rightarrow z \;\Rightarrow\; y \overset{\sim}{\downarrow} z$$

Lemma 2.1.12 *Für jede* terminierende *Relation* $\rightarrow$ *und jede Äquivalenzrelation* $\sim$ *gilt:*
$\rightarrow$ *ist konfluent modulo* $\sim$ $\quad\Leftrightarrow\quad$ $\rightarrow$ *ist lokal konfluent modulo* $\sim$

Lemma 2.1.13 *Sei* $\vdash\!\dashv$ *eine symmetrische Relation,* $\sim = \overset{*}{\vdash\!\dashv}$ *und* $\rightarrow$ *eine Relation für die* $\rightarrow \cdot \sim$ *terminierend ist. Dann ist* $\rightarrow$ *genau dann konfluent modulo* $\sim$*, wenn die Eigenschaften* $\mathcal{A}$ *und* $\mathcal{C}$ *erfüllt sind.*

Lemma 2.1.14 *Seien* $\rightarrow_A$*,* $\rightarrow_B$ *und* $\rightarrow_C$ *Relationen auf einer Menge* M*, wobei* $\rightarrow_C$ *terminierend und* $\rightarrow_C^{=}$ *eine partielle Ordnung ist und außerdem gilt:*

$$\begin{aligned} s \rightarrow_A t &\;\Rightarrow\; s \rightarrow_C^{=} t\\ s \rightarrow_B t &\;\Rightarrow\; s \rightarrow_C t\end{aligned}$$

Dann gilt folgende Implikation:

$$\begin{aligned}&(\forall x, x', y.\ x \rightarrow_B x' \;\wedge\; x \rightarrow_A y \;\Rightarrow\; \exists y'.\ x' \rightarrow_A^* y' \;\wedge\; y \rightarrow_B^* y')\\ &\Longrightarrow\\ &(\forall x, x', y.\ x \rightarrow_B^* x' \;\wedge\; x \rightarrow_A^* y \;\Rightarrow\; \exists y'.\ x' \rightarrow_A^* y' \;\wedge\; y \rightarrow_B^* y')\end{aligned}$$

Graphisch dargestellt sieht dies so aus:

$$\begin{tikzcd} x \arrow[r, "A"'] \arrow[d, "B"'] & y \arrow[d, dashed, "B"', "*"] \\ x' \arrow[r, dashed, "*", "A"'] & y' \end{tikzcd} \quad\Longrightarrow\quad \begin{tikzcd} x \arrow[r, "*", "A"'] \arrow[d, "B"', "*"] & y \arrow[d, dashed, "B"', "*"] \\ x' \arrow[r, dashed, "*", "A"'] & y' \end{tikzcd}$$

Beweis durch wohlfundierte Induktion über x mit der Relation $\rightarrow_C$. $\square$

2.2 Definitionen und Lemmata zu Termersetzungssystemen

2.2.1 Kritische Paare

Um zu entscheiden, ob ein HRS konfluent ist, ist es nötig seine kritischen Paare zu untersuchen. Kritische Paare wurden zuerst von Knuth und Bendix [10] für Termersetzungssysteme erster Ordnung eingeführt und später von Nipkow [2] auf HRS verallgemeinert. Die Grundidee dabei ist, zu testen ob die linke Seite einer Regel mit einem Subterm einer linken Seite (eventuell auch mit einem Subterm von sich selbst) unifizierbar ist. Man sagt dann, daß die beiden Regeln überlappen. Wenn dies der Fall ist könnte die Reduktion eines Terms durch Anwendung dieser Regeln (manchmal an der gleichen, doch meist an verschiedenen Positionen in t) in zwei verschiedene Richtungen laufen. Da der Fall auftreten kann, daß Variablen, die in einer linken Seite l gebunden sind in einem Subterm von l frei sind, ist es für die Definition der kritischen Paare nötig, eine Hilfskonstruktion einzuführen, die die so frei gewordenen Variablen wieder bindet.

Definition 2.2.1 Für einen λ-Term t und $p \in Pos(t)$ sei $bv(t,p)$ die Menge der Variablen die auf dem Pfad von der Wurzel von t zur Position p durch λ-Abstraktion gebunden werden. Diese Menge wird induktiv über den Aufbau von t und p definiert wie folgt:

$$\begin{array}{rcl} bv(t,\epsilon) & = & \emptyset \\ bv(\lambda x.s,p) & = & \{x\} \cup bv(s, rest(p)) \\ bv((s_1\ s_2),p) & = & \begin{cases} bv(s_1, rest(p)) & \text{falls } first(p) = 1 \\ bv(s_2, rest(p)) & \text{falls } first(p) = 2 \end{cases} \end{array}$$

Die Position p ist dabei, wie schon früher definiert, eine Sequenz über $\{1,2\}$. Die Funktion $first$ liefert dabei das erste Element der Sequenz, und $rest(p)$ liefert als Ergebnis die Sequenz, die aus p entsteht, wenn das erste Element entfernt wird. Diese Definition ist vollständig, weil $p \in Pos(t)$.

Ein $\overline{x_k}$**-lifter** eines Terms t **weg von** W ist eine Substitution $\sigma = \{F \mapsto (\rho F)(\overline{x_k}) \mid F \in fv(t)\}$, wobei ρ eine Umbenennung mit $Dom(\rho) = fv(t)$, $Cod(\rho) \cap W = \emptyset$ und $\rho F : \tau_1 \to \cdots \to \tau_k \to \tau$ für $x_1 : \tau_1, \ldots, x_k : \tau_k$ und $F : \tau$ ist.

Definition 2.2.2 (Kritisches Paar) Seien zwei Regeln $l_i \to r_i$, $i = 1,2$ in einem HRS gegeben, sowie eine Position $p \in Pos(l_1)$ und ein allgemeinster Unifikator θ von $\lambda\overline{x_k}.(l_1/p)$ und $\lambda\overline{x_k}.(\sigma(l_2))$ für die gilt:

- $fv(l_1) \cap bv(l_1) = \emptyset$
- Der Kopf von l_1/p ist keine freie Variable in l_1

- $\{\overline{x_k}\} = bv(l_1, p)$ und σ ist ein $\overline{x_k}$-lifter von l_2 weg von $fv(l_1)$

Dann heißen die beiden Regeln $l_1 \to r_1$ und $l_2 \to r_2$ **überlappend** an Position p und es ist dadurch das **kritische Paar** $\langle \theta(r_1)\downarrow_\beta,\ \theta(l_1[\sigma(r_2)]_p)\downarrow_\beta \rangle$ bestimmt.

Zwei Redexe t/p_1 und t/p_2 in einem Term t heißen **kritisch** zueinander, wenn $p_1 \leq p_2$ und $t/p_1 = \theta_1(l_1)\downarrow_\beta$ und $t/p_2 = \theta_2(l_2)\downarrow_\beta$ und $(l_1 \to r_1)$ mit $(l_2 \to r_2)$ an Position p_2/p_1 überlappt.

Bemerkungen:

- Es muß nur vorausgesetzt werden, daß die Terme $\lambda\overline{x_k}.(l_1/p)$ und $\lambda\overline{x_k}.(\sigma(l_2))$ unifizierbar sind. Dann folgt die Existenz des allgemeinsten Unifikators θ aus Satz 1.2.3, da die beiden Terme nach ihrer Konstruktion Patterns sind.

- Aus der Definition folgt, daß für ein kritisches Paar $\langle s, t\rangle$ sowohl s als auch t in langer $\beta\eta$-Normalform sind.

- Ein kritisches Paar $\langle u_1, u_2\rangle$ ist ein *geordnetes* Paar. Wenn $\langle u_1, u_2\rangle$ ein kritisches Paar eines HRS R ist, so ist $\langle u_2, u_1\rangle$ nicht notwendigerweise auch ein kritisches Paar.

Die kritischen Paare eines HRS R sind alle kritischen Paare die durch Überlappung von irgendwelchen zwei Regeln aus R entstehen. Davon ausgenommen sind die trivialen Überlappungen einer Regel mit sich selbst an der Position ϵ. Es kann jedoch durchaus vorkommen, daß eine Regel mit sich selbst an einer Position $p \neq \epsilon$ überlappt und ein kritisches Paar bestimmt.

Da diese Definition in Beweisen umständlich zu handhaben ist, wird im folgenden Lemma eine etwas einfachere Charakterisierung von kritischen Paaren angegeben.

Lemma 2.2.3 *Seien zwei Regeln $l_i \to r_i,\ i = 1, 2$ und eine Position $p \in Pos(l_1)$ gegeben, wobei l_1/p nicht von der Form $F(\ldots)$ ist. Seien außerdem θ_1 und θ_2 zwei Substitutionen mit $\theta_1(l_1/p)\downarrow_\beta = \theta_2(l_2)\downarrow_\beta$ und $bv(l_1) \cap Cod(\theta_1) = \emptyset$. Dann existiert ein kritisches Paar $\langle s, t\rangle$ und eine Substitution δ so daß $\delta(s)\downarrow_\beta = \theta_1(r_1)\downarrow_\beta$ und $\delta(t)\downarrow_\beta = \theta_1(l_1)\downarrow_\beta[\theta_2(r_2)\downarrow_\beta]_p$.*

Beweis O.B.d.A sei $Dom(\theta_1) \subseteq fv(l_1)$. Sei $\{\overline{x_k}\} = bv(l_1/p)$. Zuerst wird gezeigt, daß $\lambda\overline{x_k}.l_1/p$ und $\lambda\overline{x_k}.\sigma(l_2)$ unifizierbar sind, wobei σ ein $\overline{x_k}$-lifter von l_2 weg von $fv(l_1)$ ist. Sei $\theta_2' := \{\rho(F) \mapsto \lambda\overline{x_k}.\theta_2(F) \mid F \in fv(l_2)\}$ und ρ die zu σ gehörige Umbenennung. Es wird gezeigt, daß $\theta_0 := \theta_1 \cup \theta_2'$ ein Unifikator ist. Aus $\theta_2'\sigma(F) = \theta_2'((\rho(F))(\overline{x_k}))\downarrow_\beta = \theta_2(F)$ für alle $F \in fv(l_2)$ folgt $\theta_2'(\sigma(l_2))\downarrow_\beta = \theta_2(l_2)\downarrow_\beta$. Somit gilt

$$\begin{array}{ll}
\theta_0(\lambda\overline{x_k}.(l_1/p))\downarrow_\beta = \theta_1(\lambda\overline{x_k}.(l_1/p))\downarrow_\beta & \mathcal{D}om(\theta_2') \cap fv(l_1) = \emptyset \\
= \lambda\overline{x_k}.\theta_1(l_1/p)\downarrow_\beta & bv(l_1) \cap \mathcal{C}od(\theta_1) = \emptyset \\
= \lambda\overline{x_k}.\theta_2(l_2)\downarrow_\beta & \\
= \lambda\overline{x_k}.\theta_2'(\sigma(l_2))\downarrow_\beta = \theta_2'(\lambda\overline{x_k}.\sigma(l_2))\downarrow_\beta & \mathcal{C}od(\theta_2') \cap \{\overline{x_k}\} = \emptyset \\
= \theta_0(\lambda\overline{x_k}.\sigma(l_2))\downarrow_\beta & \sigma \text{ weg von } fv(l_1) \supseteq \mathcal{D}om(\theta_1)
\end{array}$$

Da die beiden Terme unifizierbar sind gibt es nach Lemma 1.2.3 einen allgemeinsten Unifikator θ von $\lambda\overline{x_k}.l_1/p$ und $\lambda\overline{x_k}.\sigma(l_2)$ und somit ein kritisches Paar $\langle s,t\rangle$, wobei $s=\theta(r_1)\downarrow_\beta$ und $t=\theta(l_1[\sigma(r_2)]_p)\downarrow_\beta$. Da σ ein allgemeinster Unifikator ist existiert eine Substitution δ mit $\theta_0=\delta\theta$ auf den freien Variablen von l_1 und l_2. Daraus folgt

$$\delta(s)\downarrow_\beta=\theta_0(r_1)\downarrow_\beta=\theta_1(r_1)\downarrow_\beta$$

und

$$\begin{aligned}\delta(t)\downarrow_\beta &= \theta_0(l_1[\sigma(r_2)]_p)\downarrow_\beta\\ &= \theta_0(l_1)\downarrow_\beta[\theta_0(\sigma(r_2))\downarrow_\beta]_p\\ &= \theta_1(l_1)\downarrow_\beta[\theta_2'(\sigma(r_2))\downarrow_\beta]_p\\ &= \theta_1(l_1)\downarrow_\beta[\theta_2(r_2)\downarrow_\beta]_p\end{aligned}$$

□

Beispiel 2.2.4 Das Termersetzungssystem aus Beispiel 1.2.6 hat zwei kritische Paare. Zur Erinnerung noch einmal seine beiden Regeln:

$$\begin{array}{lrcl}\beta & : \; app(abs(\lambda x.F(x)),S) & \to & F(S)\\ \eta & : \; abs(\lambda x.app(H,x)) & \to & H\end{array}$$

Das erste kritische Paar ergibt sich aus der Überlappung der linken Seite der η-Regel mit einem Subterm der linken Seite der β-Regel. Dabei wird für die freie Variable F der Term $\lambda y.app(H,y)$ substituiert. Daraus folgt:

$$\begin{array}{lcl} & & (\lambda y.app(H,y)\ S)\downarrow_\beta \;=\; app(H,S)\\ & \nearrow_\beta & \\ app(abs(\lambda x.app(H,x)),S) & & \\ & \searrow_\eta & \\ & & app(H,S)\end{array}$$

Das erste kritische Paar ist also $\langle app(H,S),app(H,S)\rangle$.

Das zweite kritische Paar ergibt sich aus der Überlappung der linken Seite der β-Regel mit einem Subterm der linken Seite der η-Regel. Dabei wird für die freie Variable H der Term $abs(\lambda y.F(y))$ und für die freie Variable S der Term x substituiert. Daraus folgt:

$$\begin{array}{lcl} & & abs(\lambda x.F(x))\\ & \nearrow_\beta & \\ abs(\lambda x.app(abs(\lambda y.F(y)),x)) & & \\ & \searrow_\eta & \\ & & abs(\lambda y.F(y))\end{array}$$

Da nach Konvention α-äquivalente Terme identifiziert werden ergibt sich das kritische Paar $\langle abs(\lambda x.F(x)),abs(\lambda x.F(x))\rangle$.

In diesem speziellen Beispiel sind beide Seiten der kritischen Paare jeweils gleich. Im allgemeinen ist dies jedoch nicht der Fall.

2.2.2 Parallele bottom-up Reduktion

Wie in Definition 1.2.4 dargestellt induziert ein HRS R eine Relation $\rightarrow$ auf der Menge der λ-Terme in langer $\beta\eta$-Normalform. Es kommt vor, daß mehrere Anwendungen von Termersetzungsregeln auf einen Term sich gegenseitig nicht wesentlich beeinflussen. Damit ist gemeint, daß die Anwendung einer dieser Regeln die Anwendbarkeit der anderen Regeln nicht beeinflußt und das Ergebnis der Reduktionen von der Reihenfolge der Regelanwendungen unabhängig ist. Das ist z.B. der Fall, wenn die Positionen der Redexe in der Baumrepräsentation des zu reduzierenden Terms alle in parallelen Unterbäumen liegen. Allgemeiner ist dies der Fall, wenn die verschiedenen Regelanwendungen nicht überlappend sind.

Beispiel 2.2.5 Sei R das HRS mit folgenden Regeln:

$$\begin{array}{rcl} f(g(F)) & \rightarrow & h(F) \\ i(F) & \rightarrow & k(l(F), m(F)) \end{array}$$

Auf den Term $\lambda x.n(f(g(c)), i(x))$ sind beide Regeln anwendbar. Da die Redexpositionen jedoch in parallelen Unterbäumen liegen beeinflussen sich die Reduktionen gegenseitig nicht. Unabhängig von der Reihenfolge der Kontraktion der Redexe $f(g(c))$ und $i(x)$ ist das Ergebnis stets $\lambda x.n(h(c), k(l(x), m(x)))$. Die Kontraktionen der Redexe kann daher in einem Schritt vorgenommen werden.

Die Redexpositionen müssen jedoch nicht unbedingt parallel sein damit die Reduktionen sich gegenseitig nicht wesentlich beeinflussen. Es genügt dazu, wenn sie nicht überlappend sind, wie das folgende Beispiel zeigt: Auf den Term $i(f(g(c)))$ sind beide Regeln anwendbar. Die Anwendung der ersten Regel ergibt

$$i(f(g(c))) \rightarrow_{1.Regel} i(h(c))$$

Die Anwendung der zweiten Regel auf diesen Term ergibt dann

$$i(h(c)) \rightarrow_{2.Regel} k(l(h(c)), m(h(c)))$$

Man kann jedoch auch die zweite Regel zuerst anwenden. Dies führt auf

$$i(f(g(c))) \rightarrow_{2.Regel} k(l(f(g(c))), m(f(g(c))))$$

Auf den so erhaltenen Term ist die erste Regel aber immer noch anwendbar. Mehrere Anwendungen der ersten Regel führen dann auf

$$\begin{array}{rcl} k(l(f(g(c))), m(f(g(c)))) & \rightarrow_{1.Regel} & k(l(h(c)), m(f(g(c)))) \\ & \rightarrow_{1.Regel} & k(l(h(c)), m(h(c))) \end{array}$$

Dies ist das gleiche Ergebnis wie im ersten Fall. Welche der Regeln zuerst angewendet wird hat also keinen Einfluß auf das Endergebnis. Um unnötige mehrfache Regelanwendungen zu vermeiden ist es sinnvoll, die Redexe deren Positionen tiefer in der Baumrepräsentation des Terms liegen zuerst zu kontrahieren.

Es liegt nahe solche Mengen von Regelanwendungen die normalerweise Schritt für Schritt nacheinander ausgeführt werden zu einem einzigen Schritt zusammenzufassen. Es wird nun eine neue durch ein HRS induzierte Relation auf den λ-Termen definiert, die solcherart gleichzeitig ausführbare Schritte zu einem Schritt zusammenfaßt. Diese Relation wird in den folgenden Kapiteln häufig benutzt werden.

Definition 2.2.6 (Parallele bottom-up Reduktion) Die durch ein HRS R induzierte parallele bottom-up Reduktion ist die kleinste Relation $>$ auf den λ-Termen, die unter den folgenden Inferenzregeln abgeschlossen ist.

$$\frac{s_i > t_i \quad (i = 1, \ldots, n)}{a(\overline{s_n}) > a(\overline{t_n})}(A) \qquad \frac{s > t}{\lambda x.s > \lambda x.t}(L) \qquad \frac{\theta >^{fv(l)} \theta'}{\theta(l)\downarrow_\beta > \theta'(r)\downarrow_\beta}(R)$$

wobei a ein Atom von Typ $\overline{\tau_n} \to \tau$, $s_i : \tau_i$, $(l \to r) \in R$ und $\theta >^W \theta'$ eine Abkürzung ist für $\forall F \in W.\ \theta(F) > \theta'(F)$.

Es folgen einige Lemmata zu den Eigenschaften dieser Reduktion:

Lemma 2.2.7 *$s > s$ gilt für alle Terme s in langer $\beta\eta$-Normalform.*
Beweis durch Induktion über den Aufbau von s. □

Lemma 2.2.8 *Wenn $s > t$ gilt, dann sind s und t in langer $\beta\eta$-Normalform.*
Beweis durch Induktion über die Struktur von $s > t$. Alle drei Inferenzregeln (A), (L), (R) erhalten die $\beta\eta$-Normalform. Für (A) gilt dies, weil $a : \overline{\tau_n} \to \tau$ und τ ein Basistyp ist. Für (L) gilt es, weil keine Applikation eines Terms auf einen Anderen vorgenommen wird, sondern nur die in s und t frei vorkommende Variable x durch λ-Abstraktion gebunden wird. Da die Terme in $Cod(\theta)$ in langer $\beta\eta$-Normalform sind ist auch $\theta(s)\downarrow_\beta$ in langer $\beta\eta$-Normalform, und somit gilt die Aussage auch für (R). □

Wie vorher angedeutet wurde soll die Reduktion $>$ mehrere Schritte der Reduktion $\to$ zu einem einzigen Schritt zusammenfassen und kann daher durch mehrere $\to$-Schritte simuliert werden.

Lemma 2.2.9 *Genau dann wenn $s > t$ gilt gibt es eine Menge $\{p_1, \ldots, p_n\} \subseteq Pos(s)$ von Positionen in s, zugehörige Regeln $(l_i \to r_i) \in R \quad (i = 1, \ldots, n)$ (die nicht alle verschieden sein müssen) so daß alle Regelanwendungen von $(l_i \to r_i)$ an p_i paarweise nicht überlappend sind und eine Bijektion $Z : \{1, \ldots, n\} \mapsto \{p_1, \ldots, p_n\}$ so daß gilt:*

$$s \to^{Z(1)} s_1 \to^{Z(2)} \cdots \to^{Z(n-1)} s_{n-1} \to^{Z(n)} t$$

wobei $\to^{p_i}$ eine Anwendung der Regel $(l_i \to r_i)$ an der Position p_i ist und folgende Bedingung gilt:

$$(*) \quad Z(a) \leq Z(b) \Rightarrow b \leq a$$

Beachte: *Das erste $\leq$ ist die partielle Präfixordnung auf der Menge $Pos(s)$ der Positionen in s, das zweite jedoch die lineare Ordnung auf den natürlichen Zahlen.*

Beweis getrennt für beide Implikationsrichtungen.

1. Es gelte
$$s \to^{Z(1)} s_1 \to^{Z(2)} \cdots \to^{Z(n-1)} s_{n-1} \to^{Z(n)} t$$
mit
$$Z(a) \leq Z(b) \Rightarrow b \leq a$$
Zu zeigen ist:
$$s > t$$
Dies wird durch vollständige Induktion über die lexikographisch geordneten Paare $(n, L(s))$ gezeigt. Dabei ist n die Anzahl der Regelanwendungen im Übergang von s nach t und $L(s)$ die im ersten Kapitel definierte Länge des λ-Terms s.

   Induktionsanfang: Sei $(n = 0, L(s) = \mathit{beliebig})$, dann gilt $s = t$. Somit gilt $s > t$, da $>$ nach Lemma 2.2.7 reflexiv ist.

   Induktionsschluß: Hier werden zwei Fälle unterschieden:

   (a) Die n-te Regelanwendung findet nicht an der Wurzel von s_{n-1} statt. Daher kann s_{n-1} kein Atom sein. Da alle Regelanwendungen nicht überlappend sind gibt es für die Gestalt von s nur zwei Möglichkeiten.

      i. $s = \lambda x.u$. Daraus folgt $t = \lambda x.u'$ mit
      $$u \to^{rest(Z(1))} \cdots \to^{rest(Z(n))} u'$$
      Dabei sind die Regelanwendungen nicht überlappend und die Bedingung $(*)$ ist erfüllt, weil dies nach Voraussetzung für den Übergang von s nach t erfüllt war.
      Aus der Definition der Länge eines Terms folgt außerdem:
      $$L(u) < L(s)$$
      Da die Anzahl der Regelanwendungen gleich geblieben ist gilt
      $$(n, L(u)) < (n, L(s))$$
      und die Induktionsvoraussetzung ist anwendbar. Somit gilt
      $$u > u'$$
      woraus mit der Inferenzregel (L) folgt:
      $$s = \lambda x.u > \lambda x.u' = t$$

ii. $s = (u_1\ u_2)$. Daraus folgt $t = (u_1{}'\ u_2{}')$ mit

$$\begin{array}{ccccc} u_1 & \to & \cdots & \to & u_1{}' \\ u_2 & \to & \cdots & \to & u_2{}' \end{array}$$

Die Regelanwendungen im Übergang von u_i zu $u_i{}'$, $(i = 1, 2)$ sind ebenfalls nicht überlappend und die Bedingung $(*)$ ist erfüllt, da dies nach Voraussetzung für den Übergang von s nach t gilt. Außerdem folgt aus der Definition der Länge von Termen:

$$L(u_i) < L(s) \quad (i = 1, 2)$$

Da die Anzahl der Regelanwendungen im Übergang von u_i nach $u_i{}'$ kleiner oder gleich n ist kann man die Induktionsvoraussetzung anwenden, woraus folgt

$$\begin{array}{ccc} u_1 & > & u_1{}' \\ u_2 & > & u_2{}' \end{array}$$

Mit der Inferenzregel (A) folgt daraus

$$s = (u_1\ u_2) > (u_1{}'\ u_2{}') = t$$

(b) Die n-te Regelanwendung finde jetzt an der Wurzel von s_{n-1} statt (wegen der Bedingung $(*)$ kann eine Regelanwendung an der Wurzel nur ganz zuletzt kommen). Es gibt also eine Regel $(l \to r)$ und eine Substitution θ' mit

$$s_{n-1} = \theta'(l)\downarrow_\beta \to \theta'(r)\downarrow_\beta = t$$

Da nach Voraussetzung alle Regelanwendungen im Übergang von s zu t nicht überlappend sind gibt es eine Substitution θ mit

$$s = \theta(l)\downarrow_\beta$$

Sei $F \in fv(l)$. Nach den Voraussetzungen über die Reduktion von s nach t und da l ein Pattern ist geht $\theta(F)$ mit *weniger* als n Reduktionsschritten über in $\theta'(F)$, wobei die Regelanwendungen nicht überlappend sind und die Bedingung $(*)$ erfüllt ist. Somit ist die Induktionsvoraussetzung anwendbar und daraus folgt

$$\forall F \in fv(l).\ \theta(F) > \theta'(F)$$

Mit der Inferenzregel (R) folgt daraus

$$s = \theta(l)\downarrow_\beta > \theta'(r)\downarrow_\beta = t$$

Damit ist die erste Implikation gezeigt.

2. Gegeben sei jetzt $s > t$. Zu zeigen ist, daß s durch eine Folge von nicht überlappenden Reduktionsschritten in t überführt werden kann, welche die Bedingung $(*)$ erfüllen. Dies wird durch Induktion über die Struktur von $s > t$ gezeigt.

   (a) s ist ein Atom. In diesem Fall wird *höchstens ein* Reduktionsschritt $s \to t$ ausgeführt. Die Bedingungen sind trivialerweise erfüllt.

   (b) $s = (s_1\ s_2) > (t_1\ t_2) = t$ mit $s_1 > t_1$ und $s_2 > t_2$. Nach Induktionsvoraussetzung gilt
   $$s_i \to \cdots \to t_i$$
   durch eine Folge von nicht überlappenden Reduktionsschritten die die Bedingung $(*)$ erfüllen. Somit gilt auch
   $$(s_1\ s_2) \to \cdots \to (t_1\ t_2)$$
   Da die Positionen in s_1 parallel zu denen in s_2 sind ist die Bedingung $(*)$ erfüllt und die Regelanwendungen sind nicht überlappend.

   (c) $s = \lambda x.u > \lambda x.u' = t$ mit $u > u'$. Aus der Induktionsvoraussetzung folgt
   $$u \to \cdots \to u'$$
   durch eine Folge von nicht überlappenden Reduktionsschritten die die Bedingung $(*)$ erfüllen. Somit gilt auch
   $$\lambda x.u \to \cdots \to \lambda x.u'$$
   Die Redexpositionen sind dabei die gleichen wie im Übergang von u nach u', nur wird noch das Präfix 1 angehängt. Somit sind auch hier die Regelanwendungen nicht überlappend und die Bedingung $(*)$ ist erfüllt.

   (d) $s = \theta(l)\downarrow_\beta > \theta'(r)\downarrow_\beta = t$ mit $\forall F \in fv(l).\ \theta(F) > \theta'(F)$. Aus der Induktionsvoraussetzung folgt
   $$\forall F \in fv(l).\ \theta(F) \to \cdots \to \theta'(F)$$
   durch nicht überlappende Reduktionen, die die Bedingung $(*)$ erfüllen. Da l ein Pattern ist, und freie Variablen daher nicht geschachtelt auftreten, folgt daraus
   $$\theta(l)\downarrow_\beta \to \cdots \to \theta'(l)\downarrow_\beta$$

durch eine Folge nicht überlappender Reduktionen, die die Bedingung $(*)$ erfüllen. Somit entspricht $\theta'(l)\downarrow_\beta$ dem Term s_{n-1}. Der letzte Reduktionsschritt

$$\theta'(l)\downarrow_\beta \to \theta'(r)\downarrow_\beta$$

findet an der Wurzel des Terms statt und ist unkritisch zu allen anderen. Somit sind alle Reduktionsschritte im Übergang von $\theta(l)\downarrow_\beta$ nach $\theta'(r)\downarrow_\beta$ zueinander unkritisch und die Bedingung $(*)$ ist erfüllt.

Damit ist auch die zweite Implikation gezeigt.

□

Man schreibt dann auch $s >^{\{p_1,\ldots,p_n\}} t$ um die Positionen der reduzierten Redexe in s anzudeuten.

Ein Beispiel für mögliche Positionen von Redexen die in einem Schritt durch die parallele bottom-up Reduktion kontrahiert werden können wird in Abbildung 2.1 gezeigt.

Korollar 2.2.10 *Sei R ein HRS, $\to$ die von R induzierte Termersetzungsrelation und $>$ die von R induzierte parallele bottom-up Reduktion. Dann gilt*

$$\to \subseteq > \subseteq \xrightarrow{*}$$

Beweis folgt aus Lemma 2.2.9. Die erste Inklusion folgt aus dem Spezialfall $n = 1$, die zweite Inklusion gilt, weil $\to^n \subseteq \to^*$ nach Definition 2.1.1. □

Lemma 2.2.11

$$s > t \ \wedge\ \theta >^{fv(s)} \theta' \quad \Rightarrow \quad \theta(s)\downarrow_\beta \ > \ \theta'(t)\downarrow_\beta$$

Beweis durch geschachtelte Induktion über die **Ordnung** von θ und die Struktur von $s > t$, wobei die Ordnung definiert ist als

$$\begin{aligned} ord(\tau) &= 0 \quad \text{falls } \tau \text{ ein Basistyp ist} \\ ord(\tau_0 \to \tau_1) &= max\{ord(\tau_0) + 1, ord(\tau_1)\} \\ ord(\theta) &= max\{ord(\tau) \mid V_\tau \cap Dom(\theta) \neq \emptyset\} \end{aligned}$$

Gemäß der Definition von $s > t$ werden drei Fälle unterschieden.

1. Wenn $s = a(\overline{s_n})$, $t = a(\overline{t_n})$ und $s_i > t_i$ für $i = 1, \ldots, n$ dann ist nach der Induktionshypothese der inneren Induktion (über $s > t$) $\theta(s_i)\downarrow_\beta \ > \ \theta'(t_i)\downarrow_\beta$.

   Falls a eine Konstante, oder eine Variable $\notin Dom(\theta)$ ist, dann gilt auch $a \notin Dom(\theta')$ und $\theta(s)\downarrow_\beta \ = \ a(\overline{\theta(s_n)\downarrow_\beta}) \ > \ a(\overline{\theta'(t_n)\downarrow_\beta}) \ = \ \theta'(t)\downarrow_\beta$.

   Sonst ist $a \in Dom(\theta)$. Weil s und t nach Lemma 2.2.8 in langer $\beta\eta$-Normalform sind gilt $\theta(a) = \lambda\overline{x_n}.u$. Nach Voraussetzung muß gelten

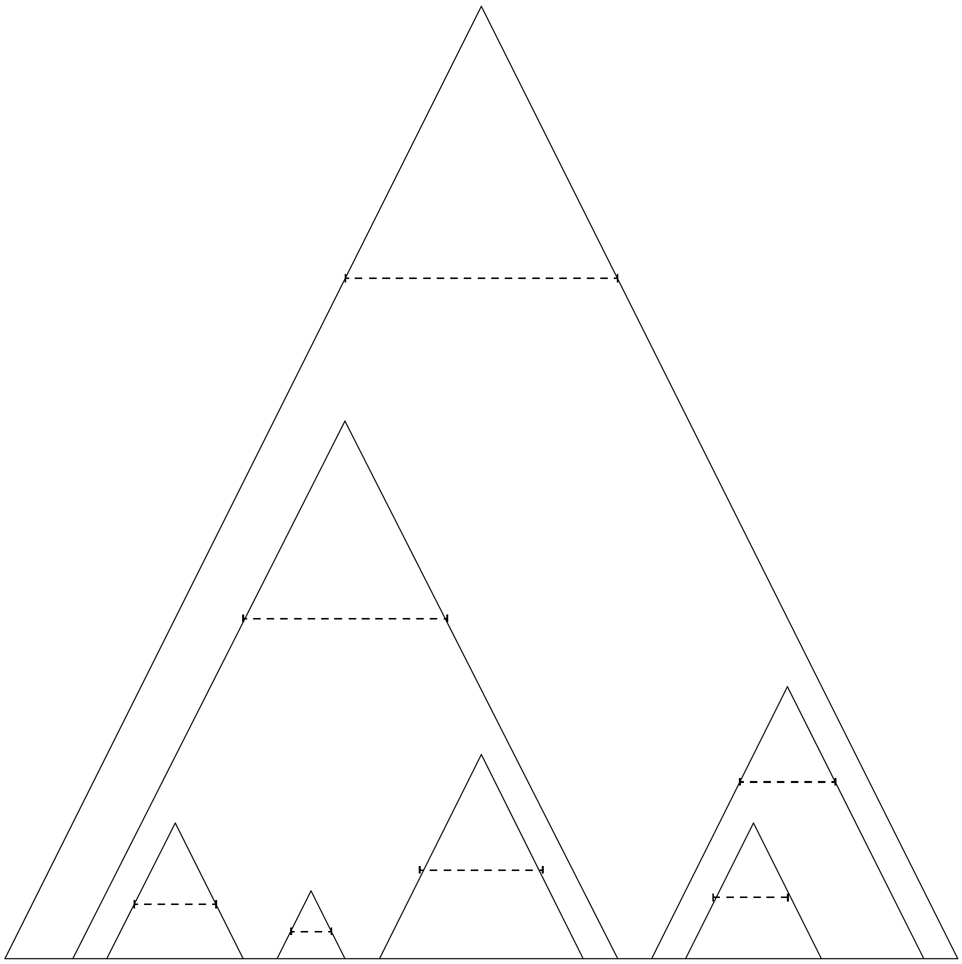


Abbildung 2.1: Ein Beispiel für einen Term mit Redexen die zueinander nicht kritisch sind. Sie können daher von unten nach oben durch die parallele bottom-up Reduktion kontrahiert werden.
Dabei ist der Teil des Redexes über der gestrichelten Linie durch die Termersetzungsregel festgelegt. Der Teil darunter entspricht den durch die Substitution in die freien Variablen der linken Seite der Regel eingesetzten Termen. Ein Redex ist somit kritisch, bzw. unkritisch zu dem darüberliegenden Redex wenn er über, bzw. unter der gestrichelten Linie des darüberliegenden Redexes beginnt.

$\theta(a) > \theta'(a)$. Da die Inferenzregel (A) aufgrund der Struktur von $\theta(a)$ nicht anwendbar ist und die Regel (R) nicht anwendbar ist, weil alle Regeln $(l \to r)$ von Basistyp sein müssen, kann der Übergang nur durch Regel (L) erfolgen. Somit ist $\theta'(a) = \lambda\overline{x_n}.u'$ mit $u > u'$. Dann gilt

$$\theta(s)\downarrow_\beta = ((\theta(a))(\overline{\theta(s_n)\downarrow_\beta}))\downarrow_\beta = \theta_1(u)\downarrow_\beta$$

und

$$\theta'(t)\downarrow_\beta = ((\theta'(a))(\overline{\theta'(t_n)\downarrow_\beta}))\downarrow_\beta = \theta'_1(u')\downarrow_\beta$$

wobei $\theta_1 = \{\overline{x_n \mapsto \theta(s_n)\downarrow_\beta}\}$ und $\theta'_1 = \{\overline{x_n \mapsto \theta'(t_n)\downarrow_\beta}\}$. Nach Definition der Form $a(\overline{s_n})$ hat a einen Typ $\overline{\tau_n} \to \tau$. Aus der Definition von ord folgt $ord(\tau_i) < ord(\overline{\tau_n} \to \tau)$ und somit $ord(\theta_1) < ord(\theta)$. Deswegen ist die Induktionshypothese der äußeren Induktion (über $ord(\theta)$) anwendbar und es gilt $\theta(s)\downarrow_\beta = \theta_1(u)\downarrow_\beta > \theta'_1(u')\downarrow_\beta = \theta'(t)\downarrow_\beta$.

2. Wenn $s = \lambda x.u$, $t = \lambda x.u'$ und $u > u'$, dann folgt aus der Induktionshypothese der inneren Induktion

$$\theta(s)\downarrow_\beta = \theta(\lambda x.u)\downarrow_\beta = \lambda x.(\theta(u)\downarrow_\beta) > \lambda x.(\theta'(u')\downarrow_\beta) = \theta'(t)\downarrow_\beta$$

3. Wenn $s = \theta_1(l)\downarrow_\beta$, $t = \theta'_1(r)\downarrow_\beta$ mit $\theta_1 >^{fv(l)} \theta'_1$, dann folgt aus der Induktionshypothese der inneren Induktion:

$$\forall F \in fv(l).\ \theta(\theta_1(F))\downarrow_\beta > \theta'(\theta'_1(F))\downarrow_\beta$$

Seien die Substitutionen θ_2 und θ'_2 definiert als

$$\begin{aligned} \theta_2 &:= \{F \mapsto \theta(\theta_1(F))\downarrow_\beta \mid F \in fv(l)\} \\ \theta'_2 &:= \{F \mapsto \theta'(\theta'_1(F))\downarrow_\beta \mid F \in fv(l)\} \end{aligned}$$

Dann gilt wegen der Induktionshypothese der äußeren Induktion

$$\begin{aligned} \theta(s)\downarrow_\beta &= \theta(\theta_1(l)\downarrow_\beta)\downarrow_\beta \\ &= \theta(\theta_1(l))\downarrow_\beta \\ &= \theta_2(l)\downarrow_\beta \\ &> \theta'_2(r)\downarrow_\beta \\ &= \theta'(\theta'_1(r))\downarrow_\beta \\ &= \theta'(\theta'_1(r)\downarrow_\beta)\downarrow_\beta \\ &= \theta'(t)\downarrow_\beta \end{aligned}$$

□

Mit Hilfe obiger Lemmata läßt es sich beweisen, daß die gerade für $>$ bewiesene Eigenschaft auch für $\to^*$ gilt.

Lemma 2.2.12 *Seien s, s' λ-Terme in langer $\beta\eta$-Normalform, θ und θ' Substitutionen mit $Dom(\theta) \subseteq Dom(\theta')$. Dann gilt*

$$\theta_{|fv(s)} \xrightarrow{*} \theta'_{|fv(s)} \wedge s \xrightarrow{*} s' \quad \Rightarrow \quad \theta(s){\downarrow_\beta} \xrightarrow{*} \theta'(s'){\downarrow_\beta}$$

Beweis Sei $M = fv(s) \cap Dom(\theta)$. Aus der Voraussetzung folgt

$$\forall F \in M.\ \theta(F) \rightarrow^* \theta'(F)$$

Da M endlich ist existiert

$i := max\{$Anzahl der Reduktionsschritte $\rightarrow$ von $\theta(F){\downarrow_\beta}$ nach $\theta'(F){\downarrow_\beta} \mid F \in M\}$

(siehe Def. 2.1.1). Daraus folgt $\theta\ (\rightarrow^=)^i\ \theta'$. Beweis der Behauptung durch geschachtelte Induktion über i und über die Länge k der Reduktion $s \rightarrow^* s'$.

Induktionsanfang der äußeren Induktion: Aus $i = 0$ folgt $\theta = \theta'$. Somit ist zu zeigen: $\theta(s){\downarrow_\beta} \rightarrow^* \theta(s'){\downarrow_\beta}$. Dies wird durch Induktion über die Länge k der Reduktion $s \rightarrow^* s'$ gezeigt.

Induktionsanfang: $k = 0 \ \Rightarrow\ s = s' \ \Rightarrow$ Beh.

Induktionsschluß: Sei $s \rightarrow^k s'$, dann gibt es nach Induktionshypothese ein s'' mit $s \rightarrow^{k-1} s'' \rightarrow s'$ und $\theta(s){\downarrow_\beta} \rightarrow^* \theta(s''){\downarrow_\beta}$. Da nach Lemma 2.2.7 $\theta > \theta$ und nach Korollar 2.2.10 $s'' > s'$ gilt folgt aus Lemma 2.2.11 $\theta(s''){\downarrow_\beta} > \theta(s'){\downarrow_\beta}$ und mit Korollar 2.2.10 auch $\theta(s''){\downarrow_\beta} \rightarrow^* \theta(s'){\downarrow_\beta}$. Mit der Transitivität von $\rightarrow^*$ folgt daraus die Behauptung.

Induktionsschluß der äußeren Induktion: Sei $\theta\ (\rightarrow^=)^i\ \theta'$, dann gibt es nach Induktionshypothese der äußeren Induktion ein θ'' mit $\theta\ (\rightarrow^=)^{i-1}\ \theta'' \rightarrow^= \theta'$ und $\theta(s){\downarrow_\beta} \rightarrow^* \theta''(s'){\downarrow_\beta}$. Aus Lemma 2.2.7 und Korollar 2.2.10 folgt $\theta'' > \theta'$ und daher nach Lemma 2.2.11 auch $\theta''(s'){\downarrow_\beta} > \theta'(s'){\downarrow_\beta}$. Mit Korollar 2.2.10 folgt $\theta''(s'){\downarrow_\beta} \rightarrow^* \theta'(s'){\downarrow_\beta}$ woraus mit der Transitivität von $\rightarrow^*$ die Behauptung folgt. □

Korollar 2.2.13 *Sei s ein λ-Term in langer $\beta\eta$-Normalform, θ und θ' Substitutionen mit $Dom(\theta) \subseteq Dom(\theta')$. Dann gilt*

$$\theta_{|fv(s)} \rightarrow^= \theta'_{|fv(s)} \quad \Rightarrow \quad \theta(s){\downarrow_\beta} \xrightarrow{*} \theta'(s){\downarrow_\beta}$$

Beweis direkt aus Lemma 2.2.12. □

Korollar 2.2.14 *Seien s und t λ-Terme in langer $\beta\eta$-Normalform und σ eine Substitution. Dann gilt*

$$s \xrightarrow{*} t \quad \Rightarrow \quad \sigma(s){\downarrow_\beta} \xrightarrow{*} \sigma(t){\downarrow_\beta}$$

Beweis direkt aus Lemma 2.2.12. □

Kapitel 3

Kriterien für lokale Konfluenz

In diesem Kapitel werden Entscheidungskriterien für die lokale Konfluenz eines HRS dargestellt. Knuth und Bendix [10] haben schon 1970 ein notwendiges und hinreichendes Kriterium für die Konfluenz von Termersetzungssystemen *erster* Ordnung (kurz: TRS) gefunden, das sich auf die Analyse der kritischen Paare stützt. Dieses Kriterium wurde von Nipkow [2] unter Verwendung einer allgemeineren Definition der kritischen Paare (siehe Def. 2.2.2) auf Termersetzungssysteme höherer Ordnung (kurz: HRS) verallgemeinert. Der Beweis dieses Ergebnisses ist jedoch in [2] nicht aufgeführt. Deswegen, und weil der Beweis viele in den Beweisen von Sätzen über Konfluenz modulo Gleichungssystemen verwendete Techniken in einfacherer Form enthält, und daher deren Verständnis erleichtert, wird er hier dargestellt.

Lemma 3.0.15 (Kritisches Paar Lemma) *Sei $\rightarrow$ die durch ein HRS R induzierte Relation auf den Termen. Wenn $s \rightarrow s_i$, $i = 1, 2$, dann gilt entweder $s_1 \downarrow s_2$, oder es gibt ein kritisches Paar $\langle u_1, u_2 \rangle$, eine Substitution δ und eine Position $p \in Pos(s)$ so daß $s_i = s[\delta(u_i)\downarrow_\beta]_p$, $i = 1, 2$.*

Beweis Nach Definition von $s \rightarrow s_i$ gibt es Regeln $(l_i \rightarrow r_i) \in R$, Positionen $p_i \in Pos(s)$ und Substitutionen θ_i mit $s/p_i = \theta_i(l_i)\downarrow_\beta$ und $s_i = s[\theta_i(r_i)\downarrow_\beta]_{p_i}$.

1. Sei $p_1 \parallel p_2$. Dann gilt
   $s_1 = s[\theta_1(r_1)\downarrow_\beta]_{p_1} \rightarrow s[\theta_1(r_1)\downarrow_\beta]_{p_1}[\theta_2(r_2)\downarrow_\beta]_{p_2} \leftarrow s[\theta_2(r_2)\downarrow_\beta]_{p_2} = s_2$.

2. Sei o.E. $p_1 \leq p_2$
   Sei $q := p_2/p_1$. Dann gilt

$$\begin{aligned} s_2/p_1 &= (s[\theta_2(r_2)\downarrow_\beta]_{p_2})/p_1 \\ &= (s/p_1)[\theta_2(r_2)\downarrow_\beta]_q \\ &= (\theta_1(l_1)\downarrow_\beta)[\theta_2(r_2)\downarrow_\beta]_q \end{aligned}$$

   und

$$\begin{aligned} (\theta_1(l_1)\downarrow_\beta)/q &= (s/p_1)/q \\ &= s/p_2 \\ &= \theta_2(l_2)\downarrow_\beta \end{aligned}$$

Es sind zwei Fälle zu unterscheiden:

(a) Die beiden Redexe überlappen nicht, d.h. $q = q_1 q_2$ so daß $q_1 \in Pos(l_1)$ und $l_1/q_1 = F(\overline{x_k})$, wobei $F \in fv(l_1)$ (Beachte: l_1 ist ein Pattern). Da nach Konvention alle Elemente von $Cod(\theta_1)$ in langer $\beta\eta$-Normalform sind gilt $\theta_1(F) = \lambda\overline{x_k}.t$. Definiere eine neue Substitution

$$\theta_1' := \theta_1 + \{F \mapsto \lambda\overline{x_k}.(t[\theta_2(r_2)\downarrow_\beta]_{q_2})\}$$

Es gilt

$$\begin{aligned} t/q_2 &= (\theta_1(F(\overline{x_k}))\downarrow_\beta)/q_2 \\ &= (\theta_1(l_1/q_1)\downarrow_\beta)/q_2 \\ &= ((\theta_1(l_1)\downarrow_\beta)/q_1)/q_2 \\ &= (\theta_1(l_1)\downarrow_\beta)/q \\ &= \theta_2(l_2)\downarrow_\beta \end{aligned}$$

und somit $\theta_1 \to^= \theta_1'$. Mit Korollar 2.2.13 folgt daraus $\theta_1(r_1)\downarrow_\beta \to^* \theta_1'(r_1)\downarrow_\beta$.

Sei H eine Variable, die bis jetzt in diesem Kontext noch nicht vorkam, sowie

$$\begin{aligned} \theta_0 &:= \theta_1 \cup \{H \mapsto \theta_1'(F)\} \\ \theta_0' &:= \theta_0 + \{F \mapsto \theta_1'(F)\} \end{aligned}$$

und $l_0 := l_1[H(\overline{x_k})]_{q_1}$. Ebenso wie $\theta_1 \to^= \theta_1'$ folgt $\theta_0 \to^= \theta_0'$, und mit Korollar 2.2.13 folgt daraus

$$\begin{aligned} (\theta_1(l_1)\downarrow_\beta)[\theta_2(r_2)\downarrow_\beta]_q &= (\theta_1(l_1)\downarrow_\beta)[\theta_1'(F(\overline{x_k}))\downarrow_\beta]_{q_1} \\ &= \theta_0(l_0)\downarrow_\beta \\ &\to^* \theta_0'(l_0)\downarrow_\beta \\ &= \theta_0'(l_1)\downarrow_\beta \\ &= \theta_1'(l_1)\downarrow_\beta \\ &\to \theta_1'(r_1)\downarrow_\beta \end{aligned}$$

Der Schritt $\theta_0(l_0)\downarrow_\beta \to^* \theta_0'(l_0)\downarrow_\beta$ ist nötig, weil die freie Variable F in l_1 mehrfach auftreten kann. Falls F nur einmal in l_1 auftritt ist $(\theta_1(l_1)\downarrow_\beta)[\theta_1'(F(\overline{x_k}))\downarrow_\beta]_{q_1} = \theta_1'(l_1)\downarrow_\beta$.

Es gilt also

$$s_1/p_1 = \theta_1(r_1)\downarrow_\beta \xrightarrow{*} \theta_1'(r_1)\downarrow_\beta \xleftarrow{*} (\theta_1(l_1)\downarrow_\beta)[\theta_2(r_2)\downarrow_\beta]_q = s_2/p_1$$

Daraus folgt

$$\begin{aligned} & & & & s_2 &= \\ & & & & s[\theta_2(r_2)\downarrow_\beta]_{p_2} &= \\ s_1 &= s[\theta_1(r_1)\downarrow_\beta]_{p_1} & \xrightarrow{*} & s[\theta_1'(r_1)\downarrow_\beta]_{p_1} & \xleftarrow{*} & s[\theta_1(l_1)\downarrow_\beta[\theta_2(r_2)\downarrow_\beta]_q]_{p_1} \end{aligned}$$

und somit $s_1 \downarrow s_2$.

(b) Die beiden Redexe überlappen, d.h. $q \in Pos(l_1)$ und l_1/q ist nicht von der Form $F(\ldots)$. O.B.d.A. sei $bv(l_1) \cap Cod(\theta_1) = \emptyset$, da dies durch α-Konversion stets erreichbar ist. Daher ist Lemma 2.2.3 anwendbar, woraus folgt, daß es ein kritisches Paar $\langle u_1, u_2 \rangle$ und eine Substitution δ gibt mit

$$\delta(u_1)\downarrow_\beta = \theta_1(r_1)\downarrow_\beta$$

und

$$\delta(u_2)\downarrow_\beta = \theta_1(l_1)\downarrow_\beta[\theta_2(r_2)\downarrow_\beta]_q$$

Somit ist

$$s_1 = s[\theta_1(r_1)\downarrow_\beta]_{p_1} = s[\delta(u_1)\downarrow_\beta]_{p_1}$$

und

$$\begin{aligned} s_2 &= s[\theta_2(r_2)\downarrow_\beta]_{p_2} \\ &= s[\theta_1(l_1)\downarrow_\beta[\theta_2(r_2)\downarrow_\beta]_q]_{p_1} \\ &= s[\delta(u_2)\downarrow_\beta]_{p_1} \end{aligned}$$

□

Mit Hilfe dieses Lemmas läßt sich ein notwendiges und hinreichendes Kriterium für die lokale Konfluenz eines HRS formulieren.

Satz 3.0.16 *Sei ein HRS R gegeben. Die von R induzierte Relation $\rightarrow$ ist genau dann lokal konfluent, wenn für alle kritischen Paare $\langle u_1, u_2 \rangle$ von R gilt: $u_1 \downarrow u_2$.*
Beweis

1. Sei $\rightarrow$ konfluent und $\langle u_1, u_2 \rangle$ ein kritisches Paar. Nach Def. 2.2.2 gibt es einen Term s mit $u_1 \leftarrow s \rightarrow u_2$. Aus der lokalen Konfluenz von $\rightarrow$ (siehe Def. 2.1.5) folgt $u_1 \downarrow u_2$.

2. Für alle kritischen Paare $\langle u_1, u_2 \rangle$ gelte $u_1 \downarrow u_2$. Sei $s \rightarrow s_i,\ i = 1, 2$. Dann ist Lemma 3.0.15 anwendbar. Da der erste Fall in diesem Lemma direkt die Behauptung ergibt bleibt nur noch der andere Fall zu untersuchen. Es gibt also ein kritisches Paar $\langle u_1, u_2 \rangle$, eine Substitution δ und eine Position $p \in Pos(s)$ so daß $s_i = s[\delta(u_i)\downarrow_\beta]_p,\ i = 1, 2$. Nach Voraussetzung gibt es ein w mit $u_1 \xrightarrow{*} w \xleftarrow{*} u_2$. Mit Korollar 2.2.14 folgt

$$\delta(u_1)\downarrow_\beta \xrightarrow{*} \delta(w)\downarrow_\beta \xleftarrow{*} \delta(u_2)\downarrow_\beta$$

und daher auch

$$s_1 = s[\delta(u_1)\downarrow_\beta]_p \xrightarrow{*} s[\delta(w)\downarrow_\beta]_p \xleftarrow{*} s[\delta(u_2)\downarrow_\beta]_p = s_2$$

Somit gilt $s_1 \downarrow s_2$ und $\rightarrow$ ist lokal konfluent.

□

Satz 3.0.17 *Die Eigenschaft der Konfluenz ist für terminierende HRS entscheidbar.*
Beweis Da ein HRS nur endlich viele Termersetzungsregeln enthält und es nur endlich viele Positionen in den linken Seiten der Regeln gibt kann es nur endlich viele kritische Paare geben. Mit Satz 1.2.3 folgt, daß diese alle berechnet werden können. Da das Termersetzungssystem terminiert ist für jedes kritische Paar $\langle u_1, u_2 \rangle$ die Eigenschaft $u_1 \downarrow u_2$ entscheidbar. Mit Satz 3.0.16 und Lemma 2.1.6 folgt daraus die Behauptung. □

Beispiel 3.0.18 Es wird noch einmal das Termersetzungssystem aus Beispiel 1.2.6 betrachtet. Wie in Beispiel 2.2.4 gezeigt wurde sind die aus Überlappungen seiner beiden Regeln

$$\begin{array}{rcrcl} \beta & : & app(abs(\lambda x.F(x)), S) & \to & F(S) \\ \eta & : & abs(\lambda x.app(H, x)) & \to & H \end{array}$$

resultierenden kritischen Paare gleich

$$\langle app(H, S), app(H, S) \rangle$$

und

$$\langle abs(\lambda x.F(x)), abs(\lambda x.F(x)) \rangle$$

Daher sind die Voraussetzungen des Satzes 3.0.16 trivialerweise erfüllt, woraus folgt, daß dieses HRS lokal konfluent ist.

Aus den bisher gezeigten Ergebnissen folgt ein hinreichendes, aber nicht notwendiges Kriterium für die Konfluenz eines HRS.

Satz 3.0.19 *Sei ein HRS R gegeben. Wenn die durch R induzierte Relation $\to$ terminierend ist und für alle kritischen Paare $\langle u_1, u_2 \rangle$ von R $u_1 \downarrow u_2$ gilt, dann ist $\to$ konfluent und jeder Term hat eine eindeutige $\to$-Normalform.*
Beweis Nach Satz 3.0.16 ist $\to$ lokal konfluent. Da die Relation $\to$ terminierend ist, ist sie nach Lemma 2.1.6 konfluent. Da $\to$ terminierend ist hat jeder Term mindestens eine $\to$-Normalform. Da $\to$ konfluent ist hat jeder Term nach Lemma 2.1.4 höchstes eine $\to$-Normalform. Somit hat jeder Term genau eine $\to$-Normalform. □

Beispiel 3.0.20 Terme und Umformungsregeln der Logik erster Stufe können durch HRS dargestellt werden. Zur Darstellung der Terme ist die Einführung der beiden Typen *term* und *form*, sowie der folgenden Konstanten nötig:

$$\begin{array}{rcl} \neg & : & form \to form \\ \wedge & : & form \to form \to form \\ \vee & : & form \to form \to form \\ \forall & : & (term \to form) \to form \\ \exists & : & (term \to form) \to form \end{array}$$

Dabei werden die Operatoren in Präfixform geschrieben. Die Normalform für die Negation, bei der der Operator $\neg$ nur auf atomare Ausdrücke angewendet wird, kann durch das HRS R mit den folgenden Regeln erreicht werden:

$$\begin{array}{rcrcl}
\neg\neg & : & \neg\neg F & \rightarrow & F \\
\neg\wedge & : & \neg\wedge(F,G) & \rightarrow & \vee(\neg F,\neg G) \\
\neg\vee & : & \neg\vee(F,G) & \rightarrow & \wedge(\neg F,\neg G) \\
\neg\forall & : & \neg\forall(\lambda x.F(x)) & \rightarrow & \exists(\lambda x.\neg F(x)) \\
\neg\exists & : & \neg\exists(\lambda x.F(x)) & \rightarrow & \forall(\lambda x.\neg F(x))
\end{array}$$

Dieses HRS hat fünf kritische Paare, die aus den Überlappungen der ersten Regel $(\neg\neg)$ mit sich selbst und den anderen vier Regeln entstehen.

1.

$$\begin{array}{lll}
 & & \neg F \\
 & \nearrow_{\neg\neg} & \\
\neg\neg\neg F & & \\
 & \searrow_{\neg\neg} & \\
 & & \neg F
\end{array}$$

Das kritische Paar $\langle \neg F, \neg F\rangle$ hat trivialerweise ein gemeinsames Redukt.

2.

$$\begin{array}{lllllll}
 & & \wedge(F,G) & & & & \\
 & \nearrow_{\neg\neg} & & & & & \\
\neg\neg\wedge(F,G) & & & & & & \\
 & \searrow_{\neg\wedge} & & & & & \\
 & & \neg\vee(\neg F,\neg G) & \rightarrow_{\neg\vee} & \wedge(\neg\neg F,\neg\neg G) & \rightarrow^*_{\neg\neg} & \wedge(F,G)
\end{array}$$

Das kritische Paar $\langle \wedge(F,G), \neg\vee(\neg F,\neg G)\rangle$ hat somit das gemeinsame Redukt $\wedge(F,G)$.

3.

$$\begin{array}{lllllll}
 & & \vee(F,G) & & & & \\
 & \nearrow_{\neg\neg} & & & & & \\
\neg\neg\vee(F,G) & & & & & & \\
 & \searrow_{\neg\vee} & & & & & \\
 & & \neg\wedge(\neg F,\neg G) & \rightarrow_{\neg\wedge} & \vee(\neg\neg F,\neg\neg G) & \rightarrow^*_{\neg\neg} & \vee(F,G)
\end{array}$$

Das kritische Paar $\langle \vee(F,G), \neg\wedge(\neg F,\neg G)\rangle$ hat somit das gemeinsame Redukt $\vee(F,G)$.

4.

$$\begin{array}{ccll}
 & & \forall(\lambda x.F(x)) & \\
 & \nearrow_{\neg\neg} & & \\
\neg\neg\forall(\lambda x.F(x)) & & & \\
 & \searrow_{\neg\forall} & & \\
 & & \neg\exists(\lambda x.\neg F(x)) & \to_{\neg\exists} \ \forall(\lambda x.\neg\neg F(x)) \\
 & & & \to_{\neg\neg} \ \forall(\lambda x.F(x))
\end{array}$$

Das kritische Paar $\langle\forall(\lambda x.F(x)), \neg\exists(\lambda x.\neg F(x))\rangle$ hat somit das gemeinsame Redukt $\forall(\lambda x.F(x))$.

5.

$$\begin{array}{ccll}
 & & \exists(\lambda x.F(x)) & \\
 & \nearrow_{\neg\neg} & & \\
\neg\neg\exists(\lambda x.F(x)) & & & \\
 & \searrow_{\neg\exists} & & \\
 & & \neg\forall(\lambda x.\neg F(x)) & \to_{\neg\forall} \ \exists(\lambda x.\neg\neg F(x)) \\
 & & & \to_{\neg\neg} \ \exists(\lambda x.F(x))
\end{array}$$

Das kritische Paar $\langle\exists(\lambda x.F(x)), \neg\forall(\lambda x.\neg F(x))\rangle$ hat somit das gemeinsame Redukt $\exists(\lambda x.F(x))$.

Da alle kritischen Paare von R ein gemeinsames Redukt haben und das System außerdem terminierend ist folgt aus Satz 3.0.19, daß R konfluent ist und die Normalformen eindeutig sind.

Das Kriterium von Satz 3.0.19 ist zwar hinreichend, aber nicht notwendig für die Konfluenz eines HRS. Es gibt auch HRS, die zwar nicht terminieren, aber dennoch konfluent sind. Dies wird in Kapitel 6 untersucht.

Kapitel 4

Konfluenz modulo Gleichungssystemen

Die Konfluenz modulo eines Gleichungssystems von Termersetzungssystemen *erster* Ordnung wurde von Huet in [1] untersucht. In diesem Kapitel werden die dort gezeigten Ergebnisse auf Termersetzungssysteme *höherer* Ordnung verallgemeinert.

4.1 Definitionen

Wie auch bei Huet [1] wird nicht die Konfluenz modulo einer beliebigen Äquivalenzrelation auf der Menge der λ-Terme untersucht, sondern es werden nur spezielle Äquivalenzrelationen betrachtet. Diese Äquivalenzrelationen müssen die Eigenschaft haben, durch ein symmetrisches HRS induziert zu sein.

Definition 4.1.1 (Induzierte Äquivalenzrelation)
Sei $\mathcal{E}$ ein **symmetrisches HRS**, d.h. für jede Regel $(l \to r) \in \mathcal{E}$ ist auch deren Umkehrung $(r \to l) \in \mathcal{E}$. Aus dieser Symmetriebedingung und Def. 1.2.4 folgt, daß für jede Regel $(l \to r) \in \mathcal{E}$ sowohl l als auch r ein Pattern sein muß. Außerdem darf weder l noch r η-äquivalent zu einer freien Variable sein und es muß gelten $fv(l) = fv(r)$. Aus der Symmetriebedingung folgt, daß die durch $\mathcal{E}$ nach Def. 1.2.4 auf den λ-Termen induzierte Relation symmetrisch ist. Sie wird daher als $\vdash\!\dashv_{\mathcal{E}}$ bzw. einfach nur als $\vdash\!\dashv$ geschrieben.

Somit ist deren reflexiv-transitive Hülle $\sim = \overset{*}{\vdash\!\dashv}$ eine Äquivalenzrelation. Sie wird als die von $\mathcal{E}$ **induzierte Äquivalenzrelation** auf den λ-Termen bezeichnet.

Definition 4.1.2 (Gleichungstheorie) Sei R ein HRS und $\mathcal{E}$ ein symmetrisches HRS. Dann heißt $\langle R, \mathcal{E}\rangle$ Gleichungstheorie.

Definition 4.1.3 (Konfluente Gleichungstheorie)
Sei R ein HRS, $\mathcal{E}$ ein symmetrisches HRS, $\rightarrow$ die durch R induzierte Relation und $\sim$ die durch $\mathcal{E}$ induzierte Äquivalenzrelation. $\langle R, \mathcal{E} \rangle$ heißt **konfluente Gleichungstheorie** genau dann, wenn $\rightarrow$ konfluent modulo $\sim$ ist.

Sei im folgenden stets R ein HRS, $\mathcal{E}$ ein symmetrisches HRS, $\rightarrow$ die von R induzierte Relation, $\vdash\!\dashv$ die von $\mathcal{E}$ induzierte symmetrische Relation und $\sim = \overset{*}{\vdash\!\dashv}$ die von $\mathcal{E}$ induzierte Äquivalenzrelation.

Definition 4.1.4 (Kritische Paare von $\mathcal{E}/R$)
Sei $\langle R, \mathcal{E} \rangle$ eine Gleichungstheorie. Die kritischen Paare von $\mathcal{E}/R$ sind definiert als alle kritischen Paare die analog zu Def. 2.2.2 aus Überlappungen von Regeln $(l_1 \rightarrow r_1)$ und $(l_2 \rightarrow r_2)$ hervorgehen, wobei entweder $(l_1 \rightarrow r_1) \in \mathcal{E}$ und $(l_2 \rightarrow r_2) \in R$ oder $(l_1 \rightarrow r_1) \in R$ und $(l_2 \rightarrow r_2) \in \mathcal{E}$ sind.

Definition 4.1.5 (Kritische Paare von $\langle R, \mathcal{E} \rangle$)
Die Menge der kritischen Paare von $\langle R, \mathcal{E} \rangle$ ist definiert als die Menge der kritischen Paare von R vereinigt mit der Menge der kritischen Paare von $\mathcal{E}/R$.

4.2 Hilfssätze

Es wird nun nach Bedingungen gesucht, unter denen ein Paar $\langle R, \mathcal{E} \rangle$ eine konfluente Gleichungstheorie ist. Aus Lemma 2.1.13 ergibt sich, daß es sinnvoll ist nach Bedingungen zu suchen, unter denen das Paar $\langle R, \mathcal{E} \rangle$, bzw. die davon induzierten Relationen, die Eigenschaften $\mathcal{A}$ und $\mathcal{C}$ aus Def. 2.1.11 erfüllen.

Lemma 4.2.1 *$\langle R, \mathcal{E} \rangle$ erfüllt Eigenschaft $\mathcal{A}$ genau dann, wenn für jedes kritische Paar $\langle u_1, u_2 \rangle$ von R gilt: $u_1 \overset{\sim}{\downarrow} u_2$.*
Beweis getrennt für beide Implikationsrichtungen.

1. Es gelte Eigenschaft $\mathcal{A}$. Sei $\langle u_1, u_2 \rangle$ ein kritisches Paar von R. Dann gibt es Regeln $(l_i \rightarrow r_i) \in R,\ i = 1, 2$, Substitutionen θ_1, θ_2 und $p \in Pos(l_1)$ mit

   $$u_1 = \theta_1(r_1)\downarrow_\beta \leftarrow \theta_1(l_1)\downarrow_\beta \rightarrow \theta_1(l_1)\downarrow_\beta[\theta_2(r_2)\downarrow_\beta]_p = u_2$$

   Somit gilt $u_1 \overset{\sim}{\downarrow} u_2$ wegen $\mathcal{A}$.

2. Für jedes kritische Paar $\langle u_1, u_2 \rangle$ von R gelte $u_1 \overset{\sim}{\downarrow} u_2$. Sei $s \rightarrow s_i,\ i = 1, 2$, dann gibt es nach Lemma 3.0.15 zwei Fälle:

   (a) $s_1 \downarrow s_2$. Mit der Reflexivität von $\sim$ folgt $s_1 \overset{\sim}{\downarrow} s_2$.

(b) Es gibt ein kritisches Paar $\langle u_1, u_2 \rangle$, eine Substitution δ und $p \in Pos(s)$ so daß $s_i = s[\delta(u_i)\downarrow_\beta]_p$, $i = 1, 2$. Nach Voraussetzung existieren Terme u'_1, u'_2 mit

$$u_1 \xrightarrow{*} u'_1 \sim u'_2 \xleftarrow{*} u_2$$

Nach Korollar 2.2.14 gilt

$$\delta(u_i)\downarrow_\beta \rightarrow^* \delta(u'_i)\downarrow_\beta, \ i = 1, 2$$

Da $\sim$ durch ein HRS induziert ist gilt auch $\delta(u'_1)\downarrow_\beta \sim \delta(u'_2)\downarrow_\beta$ nach Korollar 2.2.14. Daher gilt

$$s_1 = s[\delta(u_1)\downarrow_\beta]_p \xrightarrow{*} s[\delta(u'_1)\downarrow_\beta]_p \sim s[\delta(u'_2)\downarrow_\beta]_p \xleftarrow{*} s[\delta(u_2)\downarrow_\beta]_p = s_2$$

und somit $s_1 \overset{\sim}{\downarrow} s_2$.

Damit ist die Eigenschaft $\mathcal{A}$ gezeigt.

$\square$

Lemma 4.2.2 *Sei $\langle R, \mathcal{E} \rangle$ eine Gleichungstheorie und R linkslinear. Eigenschaft $\mathcal{C}$ ist genau dann erfüllt, wenn für jedes kritische Paar $\langle u_1, u_2 \rangle$ von $\mathcal{E}/R$ gilt: $u_1 \overset{\sim}{\downarrow} u_2$.*

Beweis getrennt für beide Implikationsrichtungen.

1. Es gelte Eigenschaft $\mathcal{C}$. Sei $\langle u_1, u_2 \rangle$ ein kritisches Paar von $\mathcal{E}/R$. Dann gibt es eine Regel $(l_1 \rightarrow r_1) \in R$, eine Regel $(l_2 \rightarrow r_2) \in \mathcal{E}$, Substitutionen θ_1, θ_2 und o.B.d.A. $p \in Pos(l_1)$ mit

$$u_1 = \theta_1(r_1)\downarrow_\beta \leftarrow \theta_1(l_1)\downarrow_\beta \vdash\!\dashv \theta_1(l_1)\downarrow_\beta[\theta_2(r_2)\downarrow_\beta]_p = u_2$$

   Somit gilt $u_1 \overset{\sim}{\downarrow} u_2$ wegen $\mathcal{C}$.

   Bemerkung: Hier wurde nur der Fall behandelt, bei dem die linke Seite einer Regel aus $\mathcal{E}$ mit einem Subterm der linken Seite einer Regel aus R überlappt. Der umgekehrte Fall geht völlig analog; es müssen nur $\rightarrow$ und $\vdash\!\dashv$ vertauscht werden.

2. Für jedes kritische Paar $\langle u_1, u_2 \rangle$ von $\mathcal{E}/R$ gelte $u_1 \overset{\sim}{\downarrow} u_2$.
   Sei $s \rightarrow s_1$ und $s \vdash\!\dashv s_2$. Nach Def. 1.2.4 gibt es Regeln $(l_1 \rightarrow r_1) \in R$, $(l_2 \rightarrow r_2) \in \mathcal{E}$, Positionen $p_1, p_2 \in Pos(s)$ und Substitutionen θ_1, θ_2 mit

$$\begin{aligned} s/p_1 &= \theta_1(l_1)\downarrow_\beta \\ s_1 &= s[\theta_1(r_1)\downarrow_\beta]_{p_1} \\ s/p_2 &= \theta_2(l_2)\downarrow_\beta \\ s_2 &= s[\theta_2(r_2)\downarrow_\beta]_{p_2} \end{aligned}$$

   Es sind drei Fälle zu unterscheiden.

(a) Sei $p_1 \parallel p_2$. Dann gilt
$s_1 = s[\theta_1(r_1)\downarrow_\beta]_{p_1} \longmapsto s[\theta_1(r_1)\downarrow_\beta]_{p_1}[\theta_2(r_2)\downarrow_\beta]_{p_2} \leftarrow s[\theta_2(r_2)\downarrow_\beta]_{p_2} = s_2$.
Somit gilt $s_1 \tilde{\downarrow} s_2$ und Eigenschaft $\mathcal{C}$ ist erfüllt.

(b) Sei $p_2 \leq p_1$ und $q := p_1/p_2$.
Dann gilt

$$\begin{aligned} s_1/p_2 &= (s[\theta_1(r_1)\downarrow_\beta]_{p_1})/p_2 \\ &= (s/p_2)[\theta_1(r_1)\downarrow_\beta]_q \\ &= (\theta_2(l_2)\downarrow_\beta)[\theta_1(r_1)\downarrow_\beta]_q \end{aligned}$$

und

$$\begin{aligned} (\theta_2(l_2)\downarrow_\beta)/q &= (s/p_2)/q \\ &= s/p_1 \\ &= \theta_1(l_1)\downarrow_\beta \end{aligned}$$

Es sind zwei Fälle zu unterscheiden:

i. Die beiden Redexe überlappen nicht, d.h. $q = q_1 q_2$ so daß $q_1 \in Pos(l_2)$ und $l_2/q_1 = F(\overline{x_k})$, wobei $F \in fv(l_2)$ (Beachte: l_2 ist ein Pattern). Da nach Konvention alle Elemente von $Cod(\theta_2)$ in langer $\beta\eta$-Normalform sind gilt $\theta_2(F) = \lambda\overline{x_k}.t$. Definiere eine neue Substitution

$$\theta_2' := \theta_2 + \{F \mapsto \lambda\overline{x_k}.(t[\theta_1(r_1)\downarrow_\beta]_{q_2})\}$$

Es gilt

$$\begin{aligned} t/q_2 &= (\theta_2(F(\overline{x_k}))\downarrow_\beta)/q_2 \\ &= (\theta_2(l_2/q_1)\downarrow_\beta)/q_2 \\ &= ((\theta_2(l_2)\downarrow_\beta)/q_1)/q_2 \\ &= (\theta_2(l_2)\downarrow_\beta)/q \\ &= \theta_1(l_1)\downarrow_\beta \end{aligned}$$

und somit $\theta_2 \rightarrow^= \theta_2'$. Mit Korollar 2.2.13 folgt daraus $\theta_2(r_2)\downarrow_\beta \rightarrow^* \theta_2'(r_2)\downarrow_\beta$.
Sei H eine Variable, die bis jetzt in diesem Kontext noch nicht vorkam, sowie

$$\begin{aligned} \theta_0 &:= \theta_2 \cup \{H \mapsto \theta_2'(F)\} \\ \theta_0' &:= \theta_0 + \{F \mapsto \theta_2'(F)\} \end{aligned}$$

und $l_0 := l_2[H(\overline{x_k})]_{q_1}$. Ebenso wie $\theta_2 \to^= \theta_2'$ folgt $\theta_0 \to^= \theta_0'$, und mit Korollar 2.2.13 folgt daraus

$$\begin{array}{rcl} & (\theta_2(l_2){\downarrow_\beta})[\theta_1(r_1){\downarrow_\beta}]_q & = & (\theta_2(l_2){\downarrow_\beta})[\theta_2'(F(\overline{x_k})){\downarrow_\beta}]_{q_1} \\ = & \theta_0(l_0){\downarrow_\beta} & \to^* & \theta_0'(l_0){\downarrow_\beta} \\ = & \theta_0'(l_2){\downarrow_\beta} & = & \theta_2'(l_2){\downarrow_\beta} \\ = & \theta_2'(l_2){\downarrow_\beta} & \vdash\!\dashv & \theta_2'(r_2){\downarrow_\beta} \end{array}$$

Der Schritt $\theta_0(l_0){\downarrow_\beta} \to^* \theta_0'(l_0){\downarrow_\beta}$ ist nötig, weil die freie Variable F in l_2 mehrfach auftreten kann. Falls F nur einmal in l_2 auftritt ist $(\theta_2(l_2){\downarrow_\beta})[\theta_2'(F(\overline{x_k})){\downarrow_\beta}]_{q_1} = \theta_2'(l_2){\downarrow_\beta}$.
Es gilt also

$$\begin{array}{ccccccc} s_2/p_2 & = & \theta_2(r_2){\downarrow_\beta} & & (\theta_2(l_2){\downarrow_\beta})[\theta_1(r_1){\downarrow_\beta}]_q & = & s_1/p_2 \\ & & \downarrow_* & & \downarrow_* & & \\ & & \theta_2'(r_2){\downarrow_\beta} & \vdash\!\dashv & \theta_2'(l_2){\downarrow_\beta} & & \end{array}$$

Daraus folgt

$$\begin{array}{ccccc} & & & s_1 & = \\ & & & s[\theta_1(r_1){\downarrow_\beta}]_{p_1} & = \\ s_2 & = & s[\theta_2(r_2){\downarrow_\beta}]_{p_2} & s[(\theta_2(l_2){\downarrow_\beta})[\theta_1(r_1){\downarrow_\beta}]_q]_{p_2} & \\ & & \downarrow_* & \downarrow_* & \\ & & s[\theta_2'(r_2){\downarrow_\beta}]_{p_2} \;\vdash\!\dashv & s[\theta_2'(l_2){\downarrow_\beta}]_{p_2} & \end{array}$$

und somit $s_1 \overset{\sim}{\downarrow} s_2$.

ii. Die beiden Redexe überlappen, d.h. $q \in Pos(l_2)$ und l_2/q ist nicht von der Form $F(\ldots)$. O.B.d.A. sei $bv(l_2) \cap Cod(\theta_2) = \emptyset$, da dies durch α-Konversion stets erreichbar ist. Daher ist Lemma 2.2.3 anwendbar, woraus folgt, daß es ein kritisches Paar $\langle u_1, u_2 \rangle$ und eine Substitution δ gibt mit $\delta(u_1){\downarrow_\beta} = \theta_2(r_2){\downarrow_\beta}$ und $\delta(u_2){\downarrow_\beta} = \theta_2(l_2){\downarrow_\beta}[\theta_1(r_1){\downarrow_\beta}]_q$. Nach Def. 4.1.4 ist $\langle u_1, u_2 \rangle$ ein kritisches Paar von $\mathcal{E}/R$. Nach Voraussetzung existieren Terme u_1', u_2' mit

$$u_1 \xrightarrow{*} u_1' \sim u_2' \xleftarrow{*} u_2$$

Nach Korollar 2.2.14 gilt

$$\delta(u_i){\downarrow_\beta} \to^* \delta(u_i'){\downarrow_\beta},\ i = 1, 2$$

Da $\sim$ durch ein HRS induziert ist gilt auch

$$\delta(u_1'){\downarrow_\beta} \sim \delta(u_2'){\downarrow_\beta}$$

nach Korollar 2.2.14. Daher gilt

$$\begin{array}{ccccccc} s_2 & = & s[\delta(u_1){\downarrow_\beta}]_{p_2} & & s[\delta(u_2){\downarrow_\beta}]_{p_2} & = & s_1 \\ & & \downarrow_* & & \downarrow_* & & \\ & & s[\delta(u_1'){\downarrow_\beta}]_{p_2} & \sim & s[\delta(u_2'){\downarrow_\beta}]_{p_2} & & \end{array}$$

und somit $s_1 \overset{\sim}{\downarrow} s_2$.

(c) Sei $p_1 \leq p_2$ und $q := p_2/p_1$.
Dann gilt

$$\begin{array}{rcl} s_2/p_1 & = & (s[\theta_2(r_2)\downarrow_\beta]_{p_2})/p_1 \\ & = & (s/p_1)[\theta_2(r_2)\downarrow_\beta]_q \\ & = & (\theta_1(l_1)\downarrow_\beta)[\theta_2(r_2)\downarrow_\beta]_q \end{array}$$

und

$$\begin{array}{rcl} (\theta_1(l_1)\downarrow_\beta)/q & = & (s/p_1)/q \\ & = & s/p_2 \\ & = & \theta_2(l_2)\downarrow_\beta \end{array}$$

Es sind zwei Fälle zu unterscheiden:

i. Die beiden Redexe überlappen nicht, d.h. $q = q_1 q_2$ so daß $q_1 \in Pos(l_1)$ und $l_1/q_1 = F(\overline{x_k})$, wobei $F \in fv(l_1)$ (Beachte: l_1 ist ein Pattern). Da nach Konvention alle Elemente von $Cod(\theta_1)$ in langer $\beta\eta$-Normalform sind gilt $\theta_1(F) = \lambda \overline{x_k}.t$. Definiere eine neue Substitution

$$\theta_1' := \theta_1 + \{F \mapsto \lambda \overline{x_k}.(t[\theta_2(r_2)\downarrow_\beta]_{q_2})\}$$

Es gilt

$$\begin{array}{rcl} t/q_2 & = & (\theta_1(F(\overline{x_k}))\downarrow_\beta)/q_2 \\ & = & (\theta_1(l_1/q_1)\downarrow_\beta)/q_2 \\ & = & ((\theta_1(l_1)\downarrow_\beta)/q_1)/q_2 \\ & = & (\theta_1(l_1)\downarrow_\beta)/q \\ & = & \theta_2(l_2)\downarrow_\beta \end{array}$$

und somit $\theta_1 \stackrel{=}{\vdash\!\dashv} \theta_1'$. Mit Korollar 2.2.13 folgt daraus $\theta_1(r_1)\downarrow_\beta \stackrel{*}{\vdash\!\dashv} \theta_1'(r_1)\downarrow_\beta$.
Da nach Voraussetzung R linkslinear ist kommt die freie Variable F in l_1 nur einmal vor. Daher gilt

$$\begin{array}{rcl} (\theta_1(l_1)\downarrow_\beta)[\theta_2(r_2)\downarrow_\beta]_q & = & (\theta_1(l_1)\downarrow_\beta)[\theta_1'(F(\overline{x_k}))\downarrow_\beta]_{q_1} \\ & = & \theta_1'(l_1)\downarrow_\beta \\ & \to & \theta_1'(r_1)\downarrow_\beta \end{array}$$

Daraus folgt

$$\begin{array}{rclcl} & & & & s_2 \quad = \\ & & & & s[\theta_2(r_2)\downarrow_\beta]_{p_2} \quad = \\ s_1 & = & s[\theta_1(r_1)\downarrow_\beta]_{p_1} & & s[(\theta_1(l_1)\downarrow_\beta)[\theta_2(r_2)\downarrow_\beta]_q]_{p_1} \\ & & \downarrow_* & & \downarrow \\ & & s[\theta_1(r_1)\downarrow_\beta]_{p_1} & \sim & s[\theta_1'(r_1)\downarrow_\beta]_{p_1} \end{array}$$

und somit $s_1 \stackrel{\sim}{\downarrow} s_2$.

ii. Die beiden Redexe überlappen, d.h. $q \in Pos(l_1)$ und l_1/q ist nicht von der Form $F(\ldots)$. O.B.d.A. sei $bv(l_1) \cap Cod(\theta_1) = \emptyset$, da dies durch α-Konversion stets erreichbar ist. Daher ist Lemma 2.2.3 anwendbar, woraus folgt, daß es ein kritisches Paar $\langle u_1, u_2 \rangle$ und eine Substitution δ gibt mit $\delta(u_1)\downarrow_\beta = \theta_1(r_1)\downarrow_\beta$ und $\delta(u_2)\downarrow_\beta = \theta_1(l_1)\downarrow_\beta[\theta_2(r_2)\downarrow_\beta]_q$. Nach Def. 4.1.4 ist $\langle u_1, u_2 \rangle$ ein kritisches Paar von $\mathcal{E}/R$. Nach Voraussetzung existieren Terme u_1', u_2' mit

$$u_1 \xrightarrow{*} u_1' \sim u_2' \xleftarrow{*} u_2$$

Nach Korollar 2.2.14 gilt

$$\delta(u_i)\downarrow_\beta \rightarrow^* \delta(u_i')\downarrow_\beta,\ i = 1, 2$$

Da $\sim$ durch ein HRS induziert ist gilt auch

$$\delta(u_1')\downarrow_\beta \sim \delta(u_2')\downarrow_\beta$$

nach Korollar 2.2.14. Daher gilt

$$\begin{array}{ccccccc} s_1 & = & s[\delta(u_1)\downarrow_\beta]_{p_1} & & s[\delta(u_2)\downarrow_\beta]_{p_1} & = & s_2 \\ & & \downarrow_* & & \downarrow_* & & \\ & & s[\delta(u_1')\downarrow_\beta]_{p_1} & \sim & s[\delta(u_2')\downarrow_\beta]_{p_1} & & \end{array}$$

und somit $s_1 \overset{\sim}{\downarrow} s_2$.

Damit ist Eigenschaft $\mathcal{C}$ gezeigt.

□

Bemerkung: Zum Beweis der Eigenschaft $\mathcal{C}$ wird die Voraussetzung der Linkslinearität nur für das HRS R, nicht aber für das HRS $\mathcal{E}$ benötigt.

4.3 Endergebnis

Mit Hilfe dieser Lemmata läßt sich nun ein Kriterium für die Konfluenz einer Gleichungstheorie formulieren.

Satz 4.3.1 *Sei $\langle R, \mathcal{E} \rangle$ eine Gleichungstheorie, R linkslinear und $(\rightarrow \cdot \sim)$ terminierend. Es stehe $\hat{t}$ für eine beliebige $\rightarrow$-Normalform eines Terms t. Dann gilt:*

Die Gleichungstheorie $\langle R, \mathcal{E} \rangle$ ist genau dann konfluent, wenn für alle kritischen Paare $\langle u_1, u_2 \rangle$ von $\langle R, \mathcal{E} \rangle$ gilt: $\widehat{u_1} \sim \widehat{u_2}$.

Beweis Aus der Reflexivität von $\sim$ folgt $\rightarrow \subseteq (\rightarrow \cdot \sim)$. Daher ist auch $\rightarrow$ terminierend und es existiert mindestens eine Normalform $\hat{t}$. Es sind zwei Implikationen zu zeigen:

1. Für alle kritischen Paare $\langle u_1, u_2 \rangle$ von $\langle R, \mathcal{E} \rangle$ gelte $\widehat{u_1} \sim \widehat{u_2}$. Daraus folgt $u_1 \overset{\sim}{\downarrow} u_2$. Mit Lemma 4.2.1 folgt daraus Eigenschaft $\mathcal{A}$. Mit der Linkslinearität von R und Lemma 4.2.2 folgt auch Eigenschaft $\mathcal{C}$. Da $(\rightarrow \cdot \sim)$ terminierend ist folgt mit Lemma 2.1.13, daß $\rightarrow$ konfluent modulo $\sim$ ist. Somit ist $\langle R, \mathcal{E} \rangle$ eine konfluente Gleichungstheorie.

2. Sei $\langle R, \mathcal{E} \rangle$ eine konfluente Gleichungstheorie, d.h. die Relation $\rightarrow$ ist konfluent modulo $\sim$. Aus Lemma 2.1.13 folgt, daß die Eigenschaften $\mathcal{A}$ und $\mathcal{C}$ erfüllt sind. Sei $\langle u_1, u_2 \rangle$ ein kritisches Paar von $\langle R, \mathcal{E} \rangle$. Dann gibt es zwei Fälle:

   (a) $\langle u_1, u_2 \rangle$ ist ein kritisches Paar von R. Dann gibt es Regeln $(l_i \rightarrow r_i) \in R,\ i = 1, 2$, Substitutionen θ_1, θ_2 und $p \in Pos(l_1)$ mit

   $$u_1 = \theta_1(r_1)\downarrow_\beta \leftarrow \theta_1(l_1)\downarrow_\beta \rightarrow \theta_1(l_1)\downarrow_\beta[\theta_2(r_2)\downarrow_\beta]_p = u_2$$

   Somit gilt $u_1 \overset{\sim}{\downarrow} u_2$ wegen $\mathcal{A}$.

   (b) $\langle u_1, u_2 \rangle$ ist ein kritisches Paar von $\mathcal{E}/R$. Dann gibt es eine Regel $(l_1 \rightarrow r_1) \in R$, eine Regel $(l_2 \rightarrow r_2) \in \mathcal{E}$, Substitutionen θ_1, θ_2 und o.B.d.A. $p \in Pos(l_1)$ mit

   $$u_1 = \theta_1(r_1)\downarrow_\beta \leftarrow \theta_1(l_1)\downarrow_\beta \vdash\!\dashv \theta_1(l_1)\downarrow_\beta[\theta_2(r_2)\downarrow_\beta]_p = u_2$$

   Somit gilt $u_1 \overset{\sim}{\downarrow} u_2$ wegen $\mathcal{C}$.

   Bemerkung: Der umgekehrte Fall, bei dem l_1 mit einem Subterm von l_2 überlappt geht völlig analog, mit $\rightarrow$ und $\vdash\!\dashv$ vertauscht.

$\square$

Kapitel 5

Gleichheit modulo Termersetzungssystemen

Termersetzungssysteme induzieren nicht nur eine operational definierte Relation (die in Abschnitt 1.2 definierte Einschrittreduktion) auf der Menge der λ-Terme, sondern auch eine logisch definierte Äquivalenzrelation. Die Eigenschaften dieser Äquivalenzrelation und ihre Beziehungen zur gewöhnlichen Reduktion werden in diesem Kapitel untersucht.

5.1 Termersetzung und Gleichheit modulo eines HRS

In diesem Abschnitt wird die Gleichheit modulo eines HRS definiert und gezeigt, daß sie modulo β- und η-Konversion mit der symmetrisch-reflexiv-transitiven Hülle der Reduktionsrelation übereinstimmt. Die meisten der folgenden Lemmata und Beweise sind in etwas anderer Form in [2] zu finden.

Definition 5.1.1 (Gleichheit modulo eines HRS) Sei ein HRS R gegeben. Die Gleichheit modulo R wird in zwei Schritten definiert. Zunächst werden die α-, β- und η-Konversion und alle Instanzen von R, d.h.

$$\{\theta(l)\downarrow_\beta = \theta(r)\downarrow_\beta \mid (l \to r) \in R\}$$

als Axiome genommen. Danach werden diese Axiome abgeschlossen unter der Reflexivität, Symmetrie, Transitivität und den beiden Kongruenzregeln

$$\frac{s = t}{\lambda x.s = \lambda x.t} \qquad \frac{s_1 = t_1 \quad s_2 = t_2}{(s_1\ s_2) = (t_1\ t_2)}$$

Die daraus resultierende Äquivalenzrelation auf der Menge der λ-Terme (die nicht notwendigerweise in langer $\beta\eta$-Normalform sein müssen) wird mit $=_R$ bezeichnet.

Das Ziel ist nun, zu zeigen, daß $s =_R t$ äquivalent ist mit $s\updownarrow_\beta^\eta \overset{*}{\underset{R}{\leftrightarrow}} t\updownarrow_\beta^\eta$. Dazu wird eine neue Relation auf der Menge der λ-Terme definiert.

Definition 5.1.2 Sei R ein HRS. Dann ist die Relation $\rightarrow_{[R]}$ definiert durch

$$s \rightarrow_{[R]} t \;:\Leftrightarrow\; \exists (l \rightarrow r) \in R, p \in Pos(s), \theta.\ (s/p)\updownarrow_\beta^\eta = \theta(l)\downarrow_\beta \;\wedge\; t = s[\theta(r)\downarrow_\beta]_p$$

Dabei ist es wichtig, zu bemerken, daß die Terme s und t *nicht* in langer $\beta\eta$-Normalform sein müssen.

Es folgen einige Lemmata über die Beziehungen zwischen der neuen Relation $\rightarrow_{[R]}$ und der normalen Reduktion $\rightarrow_R$.

Lemma 5.1.3 *Seien s und t λ-Terme in langer $\beta\eta$-Normalform. Dann gilt*

$$s \rightarrow_R t \;\Rightarrow\; s \rightarrow_{[R]} t$$

Beweis direkt aus der Definition von $\rightarrow_R$ und $\rightarrow_{[R]}$, weil $s/p = (s/p)\updownarrow_\beta^\eta$. □

Lemma 5.1.4 *Die Relation $\rightarrow_{[R]}$ erhält die $\beta\eta$-Normalform, d.h.*

$$s\updownarrow_\beta^\eta \rightarrow_{[R]} t \;\Rightarrow\; s\updownarrow_\beta^\eta \rightarrow_R t\updownarrow_\beta^\eta$$

Beweis t ist in $\beta\eta$-Normalform, weil alle Regeln in R von Basistyp sind. □

Korollar 5.1.5 *Es gilt*

$$s\updownarrow_\beta^\eta \rightarrow_{[R]}^* t \;\Rightarrow\; s\updownarrow_\beta^\eta \rightarrow_R^* t\updownarrow_\beta^\eta$$

Beweis direkt aus Lemma 5.1.4. □

Lemma 5.1.6 *Sei ein HRS R gegeben. Dann gilt*

$$s \rightarrow_{[R]} t \;\wedge\; s \rightarrow_\beta s' \;\Rightarrow\; \exists t'.\ s' \rightarrow_{[R]}^* t' \;\wedge\; t \rightarrow_\beta^* t'$$

Beweis Sei B der im Übergang von s nach s' kontrahierte β-Redex. Er hat die Form

$$B = ((\lambda x.w)\ u)$$

Sein Kontraktum ist der Term $w[x := u]$, bei dem der Term u für die Variable x in w eingesetzt worden ist. Sei außerdem D der im Übergang von s nach t kontrahierte $\rightarrow_{[R]}$-Redex und D' sein Kontraktum. Sowohl D als auch B sind Subterme von s. Abhängig von der Positionen der Redexe B und D sind drei Fälle zu unterscheiden.

1. D und B liegen in parallelen Unterbäumen von s. Ihre Kontraktionen beeinflussen sich daher gegenseitig nicht. Somit gilt

$$\begin{array}{ccc} s & \rightarrow_{[R]} & t \\ \downarrow_\beta & & \downarrow_\beta \\ s' & \rightarrow_{[R]} & t' \end{array}$$

2. B ist Subterm von D. Daraus folgt $t = t'$ und somit

$$\begin{array}{ccc} s & \rightarrow_{[R]} & t \\ \downarrow_\beta & & \downarrow_0^\beta \\ s' & \rightarrow_{[R]} & t' \end{array}$$

3. D ist Subterm von B. Hier gibt es wieder zwei Fälle.

   (a) D ist Subterm von u. Sei u' der Term, der aus u entsteht, wenn der Redex D kontrahiert wird. Dann gilt

$$\begin{array}{ccc} ((\lambda x.w)\ u) & \rightarrow_{[R]} & ((\lambda x.w)\ u') \\ \downarrow_\beta & & \downarrow_\beta \\ w[x := u] & \rightarrow_{[R]}^* & w[x := u'] \end{array}$$

   Für den Übergang von $w[x := u]$ nach $w[x := u']$ sind i.a. mehrere $\rightarrow_{[R]}$-Schritte nötig, da die Variable x mehrfach in w auftreten kann. Sei $p \in Pos(s)$ die Position von B in s. Dann gilt $t' = s[w[x := u']]_p$ und

$$\begin{array}{ccc} s & \rightarrow_{[R]} & t \\ \downarrow_\beta & & \downarrow_\beta \\ s' & \rightarrow_{[R]}^* & t' \end{array}$$

   (b) In diesem letzten Fall ist D ein Subterm von w an der Position $q \in Pos(w)$, weil alle Regeln von Basistyp sind.

   Falls die Variable x nicht in D vorkommt, so beeinflussen sich die Reduktionen der Redexe B und D gegenseitig nicht, und die Behauptung gilt trivialerweise wie im 1. Fall.

   Sei $(l \rightarrow r) \in R$ die zum Redex D gehörende Termersetzungsregel, d.h. es gibt eine Substitution θ mit

$$(w/q)\updownarrow_\beta^\eta = \theta(l)\downarrow_\beta$$

   Die Variable x, die in B durch λ-Abstraktion gebunden wird ist in w eine freie Variable. Sie kann daher nicht direkt durch die Termersetzungsregel $(l \rightarrow r)$ angesprochen werden, sondern wird durch die

Substitution θ imitiert, d.h. sie ist kein Subterm von l, sondern ein Subterm der Parameter von l. Daher gibt es eine Substitution θ' mit

$$(w[x := u]/q)\updownarrow_\beta^\eta \;=\; \theta'(l)\downarrow_\beta \;=\; \theta(l)[x := u]\downarrow_\beta$$

Daraus folgt

$$\begin{array}{ccc} ((\lambda x.w)\ u) & \to_{[R]} & ((\lambda x.w[\theta(r)\downarrow_\beta]_q)\ u) \\ & & \downarrow_\beta \\ \downarrow_\beta & & (w[x := u])[\theta(r)\downarrow_\beta[x := u]]_q \\ & & \downarrow_*^\beta \\ w[x := u] & & (w[x := u])[\theta(r)[x := u]\downarrow_\beta]_q \\ = (w[x := u])[\theta'(l)\downarrow_\beta]_q & \to_{[R]} & = (w[x := u])[\theta'(r)\downarrow_\beta]_q \end{array}$$

Sei $p \in Pos(s)$ die Position von B in s. Dann gilt

$$t' = s[(w[x := u])[\theta'(r)\downarrow_\beta]_q]_p$$

und damit

$$\begin{array}{ccc} s & \to_{[R]} & t \\ \downarrow_\beta & & \downarrow_*^\beta \\ s' & \to_{[R]} & t' \end{array}$$

□

Aus dem Beweis wird klar, daß i.a. ein $\to_\beta$- *bzw.* $\to_{[R]}$*-Schritt nicht ausreicht, um von t bzw. s' zum gemeinsamen Redukt t' zu kommen. Die reflexiv-transitiven Hüllen der Relationen beim Übergang $t \to_\beta^* t'$ und $s' \to_{[R]}^* t'$ sind also unverzichtbar.*

Dieses Lemma läßt sich noch verallgemeinern.

Lemma 5.1.7 *Sei ein HRS R gegeben. Dann gilt*

$$s \to_{[R]}^* t \;\wedge\; s \to_\beta^* s' \;\Rightarrow\; \exists t'.\ s' \to_{[R]}^* t' \;\wedge\; t \to_\beta^* t'$$

Beweis Sei $[t]_\beta$ die Länge der längsten Kette von β-Reduktionen, die mit dem Term t anfangen. Dann wird die Relation $\to_C$ auf der Menge der λ-Terme definiert durch

$$s \to_C t \quad :\Leftrightarrow \quad [t]_\beta < [s]_\beta$$

Da die β-Reduktion im einfach typisierten λ-Kalkül terminierend ist, ist $\to_C$ terminierend. Die Relation $\to_C^=$ ist trivialerweise reflexiv und transitiv.

Aus $s \to_C^= t \;\wedge\; t \to_C^= s$ folgt $([t]_\beta < [s]_\beta \;\vee\; s = t) \wedge ([s]_\beta < [t]_\beta \;\vee\; t = s)$ und daraus folgt $s = t$. Daher ist $\to_C^=$ auch antisymmetrisch und somit eine partielle Ordnung.

Außerdem gilt:

$$\begin{array}{ccc} s \to_\beta t & \Rightarrow & s \to_C t \\ s \to_{[R]} t & \Rightarrow & s \to_C^= t \end{array}$$

Damit folgt die Behauptung aus Lemma 5.1.6 und aus dem relationentheoretischen Lemma 2.1.14. □

Damit läßt sich zeigen, daß die Relation $\rightarrow^*_{[R]}$ modulo $\beta\eta$-Konversion mit der Relation $\rightarrow^*_R$ übereinstimmt.

Lemma 5.1.8 *Sei R ein HRS und s und t λ-Terme (die i.a. nicht in $\beta\eta$-Normalform sind). Dann gilt*

$$s \rightarrow^*_{[R]} t \;\Rightarrow\; s\updownarrow^\eta_\beta \rightarrow^*_R t\updownarrow^\eta_\beta$$

Beweis Es gilt $s \rightarrow^*_\beta s\downarrow_\beta$. Aus Lemma 5.1.7 folgt daher, daß es ein t' gibt mit

$$s\downarrow_\beta \rightarrow^*_{[R]} t'$$

und

$$t \rightarrow^*_\beta t'$$

Da $\rightarrow_{[R]}$ die β-Normalform erhält gilt $t' = t'\downarrow_\beta$ und daher auch $t' = t\downarrow_\beta$. Daraus folgt

$$s\updownarrow^\eta_\beta \rightarrow^*_{[R]} t\updownarrow^\eta_\beta$$

Mit Korollar 5.1.5 folgt daraus die Behauptung. □

Sei $\geq_R$ die Relation auf den λ-Termen, die auf fast die gleiche Weise wie $=_R$ gebildet wird, mit dem einzigen Unterschied, daß die Symmetrie dabei weggelassen wird.

Für diese Relation wird ein weiteres Lemma gezeigt.

Lemma 5.1.9 *Sei R ein HRS und s, t λ-Terme. Dann gilt*

$$s \geq_R t \quad\Leftrightarrow\quad s\updownarrow^\eta_\beta \rightarrow^*_R t\updownarrow^\eta_\beta$$

Beweis

1. Die $\Leftarrow$-Implikation folgt einfach aus den Definitionen.

2. Für die $\Rightarrow$-Implikation wird dies durch Induktion über die Struktur von $s \geq_R t$ bewiesen. Dabei sind mehrere Fälle zu unterscheiden.

   (a) Für $\theta(l)\downarrow_\beta \geq_R \theta(r)\downarrow_\beta$ gilt $\theta(l)\downarrow_\beta \rightarrow_{[R]} \theta(r)\downarrow_\beta$. Die Behauptung folgt daraus, daß $\theta(l)\downarrow_\beta$ und $\theta(r)\downarrow_\beta$ in langer $\beta\eta$-Normalform sind.

   (b) Für die Regeln der Reflexivität und Transitivität gibt es ebenfalls keine Probleme, da auch $\rightarrow^*_R$ reflexiv und transitiv ist.

   (c) Für die α-Konversion treten keine Probleme auf, da α-äquivalente Terme nach Konvention identifiziert werden.

   (d) Für die β- und η-Konversion ist nur zu zeigen, daß aus $s\updownarrow^\eta_\beta \rightarrow^*_R t\updownarrow^\eta_\beta$ und $s =_{\beta\eta} s'$ und $t =_{\beta\eta} t'$ die Aussage $s'\updownarrow^\eta_\beta \rightarrow^*_R t'\updownarrow^\eta_\beta$ folgt. Dies gilt, weil die $\beta\eta$-Normalform im typisierten λ-Kalkül existiert und eindeutig ist, und somit $s\updownarrow^\eta_\beta = s'\updownarrow^\eta_\beta$ und $t\updownarrow^\eta_\beta = t'\updownarrow^\eta_\beta$ gilt.

(e) Für die λ-Abstraktion ist nur zu zeigen, daß mit $s\updownarrow^\eta_\beta \rightarrow_R t\updownarrow^\eta_\beta$ auch

$$(\lambda x.s)\updownarrow^\eta_\beta \rightarrow_R (\lambda x.t)\updownarrow^\eta_\beta$$

gilt. Dies folgt direkt aus $(\lambda x.s)\updownarrow^\eta_\beta = \lambda x.(s\updownarrow^\eta_\beta)$, $(\lambda x.t)\updownarrow^\eta_\beta = \lambda x.(t\updownarrow^\eta_\beta)$ und der Definition von $\rightarrow_R$.

(f) Es bleibt der Fall der Applikation zu untersuchen, wo $(s_1\ s_2) \geq_R (t_1\ t_2)$ aus $s_i \geq_R t_i$ $(i = 1, 2)$ folgt.

Es gilt also nach Induktionsvoraussetzung $s_i\updownarrow^\eta_\beta \rightarrow^*_R t_i\updownarrow^\eta_\beta$, $(i = 1, 2)$. Mit Lemma 5.1.3 folgt daraus $s_i\updownarrow^\eta_\beta \rightarrow^*_{[R]} t_i\updownarrow^\eta_\beta$. Nach Definition von $\rightarrow_{[R]}$ folgt daraus

$$(s_1\updownarrow^\eta_\beta\ s_2\updownarrow^\eta_\beta) \rightarrow^*_{[R]} (t_1\updownarrow^\eta_\beta\ t_2\updownarrow^\eta_\beta)$$

Mit Lemma 5.1.8 folgt daraus

$$(s_1\ s_2)\updownarrow^\eta_\beta = (s_1\updownarrow^\eta_\beta\ s_2\updownarrow^\eta_\beta)\updownarrow^\eta_\beta \rightarrow^*_R (t_1\updownarrow^\eta_\beta\ t_2\updownarrow^\eta_\beta)\updownarrow^\eta_\beta = (t_1\ t_2)\updownarrow^\eta_\beta$$

Damit ist auch dieser Fall gezeigt.

Dieser Fall war der Grund für die Einführung der Hilfsrelation $\rightarrow_{[R]}$, da die Aussage

$$s_i\updownarrow^\eta_\beta \rightarrow_R t_i\updownarrow^\eta_\beta\ (i = 1, 2)\ \Rightarrow\ (s_1\ s_2)\updownarrow^\eta_\beta \rightarrow_R (t_1\ t_2)\updownarrow^\eta_\beta$$

i.a. falsch ist.

Mit Hilfe der in Abschnitt 2.2.2 bewiesenen Lemmata kann für diesen Fall auch ein Beweis angegeben werden, der ohne die Hilfsrelation $\rightarrow_{[R]}$ auskommt. Nach Voraussetzung gilt $s_i\updownarrow^\eta_\beta \rightarrow^*_R t_i\updownarrow^\eta_\beta$, $(i = 1, 2)$, und zu zeigen ist $(s_1\ s_2)\updownarrow^\eta_\beta \rightarrow^*_R (t_1\ t_2)\updownarrow^\eta_\beta$. Dazu werden zunächst zwei neue Substitutionen definiert.

$$\begin{array}{lcl} \theta & := & \{x \mapsto s_2\updownarrow^\eta_\beta\} \\ \theta' & := & \{x \mapsto t_2\updownarrow^\eta_\beta\} \end{array}$$

Aus der Voraussetzung folgt direkt $\theta \rightarrow^*_R \theta'$. Aus der Definition der langen $\beta\eta$-Normalform folgt $s_1\updownarrow^\eta_\beta = \lambda x.s'_1$ und $t_1\updownarrow^\eta_\beta = \lambda x.t'_1$, wobei s'_1 und t'_1 in langer $\beta\eta$-Normalform sind. Aus der Voraussetzung folgt $s'_1 \rightarrow^*_R t'_1$. Somit gilt:

$$\begin{array}{rcll} (s_1\ s_2)\updownarrow^\eta_\beta & = & (s_1\updownarrow^\eta_\beta\ s_2\updownarrow^\eta_\beta)\updownarrow^\eta_\beta & \text{Def. von } \updownarrow^\eta_\beta \\ & = & \theta(s'_1)\updownarrow^\eta_\beta & \text{Def. von } \theta \\ & = & \theta(s'_1)\downarrow_\beta & s'_1, Cod(\theta) \text{ in } \beta\eta\text{-Nf.} \\ & \rightarrow^*_R & \theta'(t'_1)\downarrow_\beta & \text{nach Lemma 2.2.12} \\ & = & \theta'(t'_1)\updownarrow^\eta_\beta & t'_1, Cod(\theta') \text{ in } \beta\eta\text{-Nf.} \\ & = & (t_1\updownarrow^\eta_\beta\ t_2\updownarrow^\eta_\beta)\updownarrow^\eta_\beta & \text{Def. von } \theta' \\ & = & (t_1\ t_2)\updownarrow^\eta_\beta & \text{Def. von } \updownarrow^\eta_\beta \end{array}$$

Der wichtigste Punkt in diesem Beweis ist die Einführung von Substitutionen zur Anwendung von Lemma 2.2.12.

Damit ist auch diese Richtung gezeigt.

□

Mit Hilfe dieses Lemmas kann man nun den gewünschten Satz über die Beziehung zwischen $=_R$ und $\overset{*}{\underset{R}{\leftrightarrow}}$ beweisen.

Satz 5.1.10 (Äquivalenz von Termersetzung und Gleichheit) *Sei R ein HRS und s, t λ-Terme. Dann gilt*

$$s =_R t \quad \Leftrightarrow \quad s\updownarrow_\beta^\eta \overset{*}{\underset{R}{\leftrightarrow}} t\updownarrow_\beta^\eta$$

Beweis direkt aus Lemma 5.1.9. □

In den Beweisen der Lemmata die zu diesem Ergebnis hinführten wurde mehrfach die Voraussetzung verwendet, daß die Termersetzungsregeln von Basistyp sind. Ohne diese Voraussetzung gilt der obige Satz nicht mehr, wie das folgende Beispiel zeigt.

Beispiel 5.1.11 Sei R das HRS, das nur die eine Regel

$$\lambda x.c(x, F(x)) \rightarrow \lambda x.d(F(x), x)$$

enthält.

Dann gilt zwar

$$c(a, f(a)) =_R d(f(a), a)$$

aber nicht

$$c(a, f(a)) \overset{*}{\underset{R}{\leftrightarrow}} d(f(a), a)$$

Der Grund dafür liegt darin, daß $\rightarrow_R$ nur auf Termen in langer $\beta\eta$-Normalform operiert. Ansonsten könnte man mit einer β-expandierten Form $c(a, f(a)) \equiv (\lambda x.c(x, f(x)))(a)$ arbeiten. Eine solcherart definierte Erweiterung von $\rightarrow_R$ wäre jedoch möglicherweise nicht mehr entscheidbar.

5.2 Ein hinreichendes Kriterium für Gleichheit modulo eines HRS

Durch Kombination der Ergebnisse über Konfluenz modulo Gleichungssystemen und über Gleichheit modulo Termersetzungssystemen läßt sich ein hinreichendes Kriterium für die Gleichheit zweier λ-Terme s und t modulo eines HRS R beweisen. Der Grundgedanke dabei ist, das gegebene HRS R in ein symmetrisches HRS $\mathcal{E}$ und ein gerichtetes HRS D aufzuspalten und dann $\langle D, \mathcal{E}\rangle$ als konfluente Gleichungstheorie zu beweisen.

Dabei kann es vorkommen, daß eine Regel $(l \to r) \in R$, die sich von ihrer Struktur her (l und r sind Patterns und $fv(l) = fv(r)$) für die Aufnahme in $\mathcal{E}$ eignen würde nur in *einer* Richtung in R vorkommt. D.h. es gilt zwar $(l \to r) \in R$, aber $(r \to l) \notin R$. Da die von R induzierte Gleichheit $=_R$ aber nach Definition symmetrisch ist verändert die Hinzunahme der Regel $(r \to l)$ zu R die Relation $=_R$ nicht. Davon ausgehend läßt sich der Begriff der zulässigen Erweiterung von R definieren.

Definition 5.2.1 (Zulässige Erweiterung von R bez. $=_R$) Sei R ein HRS und

$$\{(l_1 \to r_1), (l_2 \to r_2), \ldots, (l_n \to r_n)\} \subseteq R$$

die Menge aller Regeln aus R für die gilt:

$$\begin{array}{l} \forall_{i \in \{1,\ldots,n\}}\ l_i \text{ und } r_i \text{ sind Patterns} \\ \forall_{i \in \{1,\ldots,n\}}\ fv(l_i) = fv(r_i) \\ \forall_{i \in \{1,\ldots,n\}}\ (r_i \to l_i) \notin R \end{array}$$

Dann ist die zulässige Erweiterung R' von R bez. $=_R$ definiert als

$$R' := R \cup \{(r_1 \to l_1), (r_2 \to l_2), \ldots, (r_n \to l_n)\}$$

Bemerkung: Die von R' induzierte Gleichheit $=_{R'}$ ist die gleiche, wie die von R induzierte Gleichheit $=_R$, da $=_R$ nach Definition symmetrisch ist.

Mit Hilfe der so definierten Erweiterung läßt sich nun das Hauptergebnis beweisen.

Satz 5.2.2 *Sei R ein HRS mit der zulässigen Erweiterung R'. Es gelte $R' = \mathcal{E} \dot\cup D$ (disjunkte Vereinigung) so daß folgende Voraussetzungen erfüllt sind:*

- *$\mathcal{E}$ ist ein symmetrisches HRS, $\vdash\!\dashv_{\mathcal{E}}$ die durch $\mathcal{E}$ induzierte Relation, $\sim := \overset{*}{\vdash\!\dashv}_{\mathcal{E}}$ und $=_{\mathcal{E}}$ die durch $\mathcal{E}$ induzierte Gleichheit.*
- *D ist ein linkslineares HRS, $\to_D$ die durch D induzierte Relation und $\hat{t}$ eine beliebige $\to_D$-Normalform eines λ-Terms t.*
- *$(\to_D \cdot \sim)$ terminiert.*
- *Für alle kritischen Paare $\langle u_1, u_2 \rangle$ der so entstandenen Gleichungstheorie $\langle D, \mathcal{E} \rangle$ gilt $\widehat{u_1} =_{\mathcal{E}} \widehat{u_2}$.*

Dann gilt folgende Äquivalenz:

$$s =_R t \quad \Leftrightarrow \quad \widehat{s\updownarrow_\beta^\eta} =_{\mathcal{E}} \widehat{t\updownarrow_\beta^\eta}$$

Beweis Aus den Voraussetzungen folgt nach Satz 4.3.1, daß $\langle D, \mathcal{E} \rangle$ eine konfluente Gleichungstheorie ist, weil die kritischen Paare in langer $\beta\eta$-Normalform sind und $=_{\mathcal{E}}$ nach Satz 5.1.10 auf der Menge der λ-Terme in langer $\beta\eta$-Normalform mit $\sim$ übereinstimmt. Da die Relation $(\rightarrow \cdot \sim)$ nach Voraussetzung terminiert und $\sim$ reflexiv ist terminiert auch $\rightarrow_D$. Daher existiert für jeden λ-Term t mindestens eine $\rightarrow_D$-Normalform $\hat{t}$.

Daher gilt

$$
\begin{array}{rcll}
s =_R t & \Leftrightarrow & s =_{R'} t & \text{Nach Definition 5.2.1} \\
 & \Leftrightarrow & s\updownarrow_\beta^\eta \overset{*}{\underset{R'}{\Leftrightarrow}} t\updownarrow_\beta^\eta & \text{Nach Satz 5.1.10} \\
 & \Leftrightarrow & s\updownarrow_\beta^\eta (\leftrightarrow_D \cup \vdash\!\dashv_{\mathcal{E}})^* \, t\updownarrow_\beta^\eta & \text{weil } R = \mathcal{E} \dot{\cup} D \\
 & \Leftrightarrow & \widehat{s\updownarrow_\beta^\eta} \sim \widehat{t\updownarrow_\beta^\eta} & \text{nach Satz 2.1.10, weil } \langle D, \mathcal{E} \rangle \text{ eine} \\
 & & & \text{konfluente Gleichungstheorie ist.} \\
 & \Leftrightarrow & \widehat{s\updownarrow_\beta^\eta} =_{\mathcal{E}} \widehat{t\updownarrow_\beta^\eta} & \text{nach Satz 5.1.10}
\end{array}
$$

$\square$

Kapitel 6

Konfluenz bei nicht-noetherschen Termersetzungssystemen

Im Satz 3.0.19 in Kapitel 3 wurde ein hinreichendes Kriterium für die Konfluenz eines HRS gezeigt. Dieses Kriterium hängt jedoch wesentlich von der Voraussetzung ab, daß die von dem HRS auf der Menge der λ-Terme induzierte Relation $\rightarrow$ terminierend bzw. noethersch ist. Da jedoch viele HRS Relationen induzieren die nicht terminierend sind, ist es sinnvoll zu fragen, unter welchen Bedingungen diese konfluent sind. Im Gegensatz zu terminierenden Relationen ist bei nicht terminierenden Relationen die Eigenschaft der lokalen Konfluenz zwar notwendig, aber nicht hinreichend für die Konfluenz. Es ist ein hinreichendes Kriterium, daß die Relation die Diamanteigenschaft besitzt, doch ist dies zu restriktiv, da nur die wenigsten durch HRS induzierten Relationen diese Eigenschaft haben. Nach Lemma 2.1.8 genügt es zum Nachweis der Konfluenz einer Relation $\rightarrow$ eine Relation $>$ zu finden, die zwischen $\rightarrow$ und der reflexiv-transitiven Hülle $\rightarrow^*$ liegt. Beide im folgenden behandelte Methoden verwenden diesen Ansatz. Sie unterscheiden sich jedoch in der Wahl der Relation $>$ und in den Bedingungen, die an das HRS, bzw. seine kritischen Paare gestellt werden.

6.1 Parallel-abgeschlossene Systeme

Im folgenden werden die Ergebnisse von Huet [1] über die Konfluenz von nicht terminierenden Termersetzungssystemen erster Ordnung auf Termersetzungssysteme höherer Ordnung (HRS) verallgemeinert. Zunächst wird eine neue durch HRS induzierte Relation eingeführt.

6.1.1 Parallele Reduktion

Definition 6.1.1 (Parallele Reduktion) Ein HRS R induziert eine Relation $\mathrel{\parallel\!\!\!\!\!\!\longrightarrow}_R$ auf der Menge der λ-Terme in langer $\beta\eta$-Normalform vermöge:

$$s \mathrel{\parallel\!\!\!\!\!\!\longrightarrow}_R t \quad :\Leftrightarrow$$

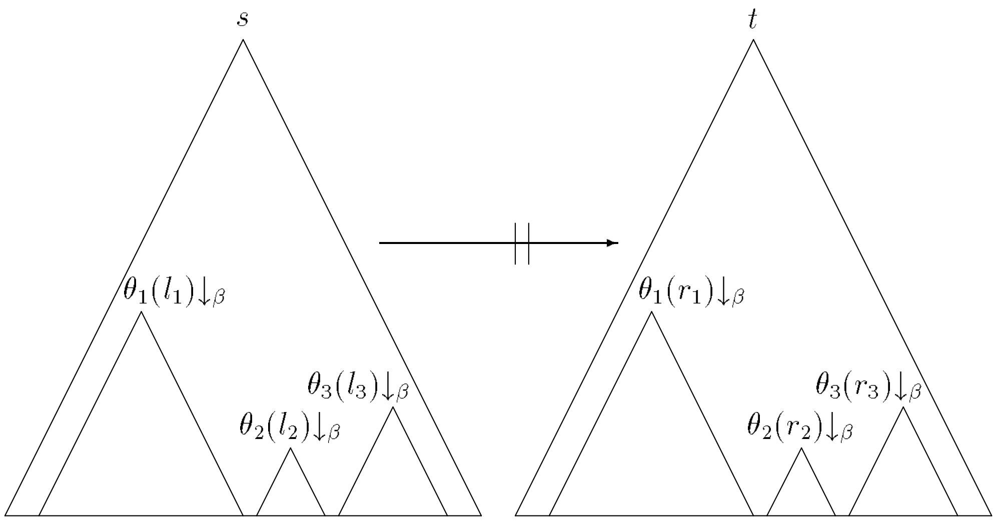
- $\exists\{p_1,\ldots,p_n\} \subseteq Pos(s)$ mit $i \neq j \;\Rightarrow\; p_i \parallel p_j$
- $\exists$ Substitutionen $\theta_1,\ldots,\theta_n$
- $\exists$ Regeln $(l_1 \to r_1),\ldots,(l_n \to r_n) \in R$

so daß gilt:

$$\forall i \in \{1,\ldots,n\}.\; s/p_i = \theta_i(l_i)\!\downarrow_\beta \quad \wedge \quad t = s[p_i \leftarrow \theta_i(r_i)\!\downarrow_\beta \;\mid\; i = 1,\ldots,n]$$

Die Relation $\mathrel{\parallel\!\!\!\!\!\!\longrightarrow}_R$ wird parallele Reduktion genannt. Wenn das HRS R in einem Kontext feststeht wird statt $\mathrel{\parallel\!\!\!\!\!\!\longrightarrow}_R$ auch nur $\mathrel{\parallel\!\!\!\!\!\!\longrightarrow}$ geschrieben. Sei $\mathcal{W} := \{p_1,\ldots,p_n\}$. Um die Menge der verwendeten Redexpositionen anzudeuten wird auch $s \mathrel{\parallel\!\!\!\!\!\!\longrightarrow}^{\mathcal{W}} t$ geschrieben. Abbildung 6.1 zeigt einen $\mathrel{\parallel\!\!\!\!\!\!\longrightarrow}$-Reduktionsschritt für $n = 3$.

Abbildung 6.1: Parallele Reduktion



Bemerkungen:

- Für ein HRS R gilt trivialerweise $\rightarrow_R \subseteq \twoheadrightarrow\!\!\!\!\!\!\shortmid\;\;_R \subseteq \rightarrow^*_R$. Zum Nachweis der ersten Inklusion wähle einfach $n = 1$. Zum Nachweis der zweiten Inklusion wende die n Regeln nacheinander an. Dies ist möglich, da alle Positionen der Regelanwendungen parallel sind und sich die Regelanwendungen daher nicht gegenseitig beeinflussen.
- Die Relation $\twoheadrightarrow\!\!\!\!\!\!\shortmid\;\;$ ist reflexiv. Für $s \twoheadrightarrow\!\!\!\!\!\!\shortmid\;\; s$ wähle einfach $n = 0$ in obiger Definition, d.h. es wird keine Regel angewendet.

6.1.2 Parallele Abgeschlossenheit

Als Verallgemeinerung der von Huet [1] eingeführten Bedingung der parallelen Abgeschlossenheit für Termersetzungssysteme erster Ordnung (TRS) wird nun die parallele Abgeschlossenheit von HRS definiert.

Definition 6.1.2 (Parallele Abgeschlossenheit) Ein HRS R heißt **parallel abgeschlossen** genau dann, wenn für jedes kritische Paar $\langle u_1, u_2 \rangle$ von R und jede Substitution σ gilt: $\sigma(u_2)\downarrow_\beta \twoheadrightarrow\!\!\!\!\!\!\shortmid\;\;_R \sigma(u_1)\downarrow_\beta$.

Bemerkung: Für Termersetzungssysteme erster Ordnung folgt die parallele Abgeschlossenheit bereits aus der Bedingung, daß für alle kritischen Paare $\langle u_1, u_2 \rangle$ von R die Bedingung

$$(PA) \quad u_2 \twoheadrightarrow\!\!\!\!\!\!\shortmid\;\; u_1$$

gilt, weil dort alle Variablen von Basistyp sind.

Für HRS ist diese Bedingung jedoch echt schwächer als die parallele Abgeschlossenheit, wie das folgende Beispiel zeigt.

Beispiel 6.1.3 Sei R das HRS mit den folgenden vier Regeln:

1. $f(\lambda y.F(y)) \rightarrow F(\lambda x.g(x))$
2. $f(\lambda y.F(y)) \rightarrow F(\lambda x.h(x))$
3. $g(F) \rightarrow h(F)$
4. $h(F) \rightarrow g(F)$

mit folgenden Typen für f, g, h :

$$\begin{array}{lcl} f & : & ((\tau_1 \rightarrow \tau_2) \rightarrow \tau_2) \rightarrow \tau_2 \\ g & : & \tau_1 \rightarrow \tau_2 \\ h & : & \tau_1 \rightarrow \tau_2 \end{array}$$

Daraus resultieren die beiden kritischen Paare

$$\langle F(\lambda x.g(x)), F(\lambda x.h(x))\rangle$$
$$\langle F(\lambda x.h(x)), F(\lambda x.g(x))\rangle$$

Für beide kritische Paare ist die Bedingung (PA) wegen

$$F(\lambda x.g(x)) \leftrightarrow_R F(\lambda x.h(x))$$

trivialerweise erfüllt.

Das HRS R ist jedoch nicht parallel abgeschlossen, da es eine Substitution θ mit

$$\theta(F) := \lambda x.x(x(c))$$

gibt mit

$$\begin{array}{rcl} \theta(F(\lambda x.g(x)))\downarrow_\beta & = & g(g(c)) \\ \theta(F(\lambda x.h(x)))\downarrow_\beta & = & h(h(c)) \end{array}$$

so daß zwar $g(g(c)) \rightarrow^* h(h(c))$, aber **nicht** $g(g(c)) \twoheadrightarrow\!\!\!\!\!\!\!\shortmid\;\; h(h(c))$ erfüllt ist.

6.1.3 Ein hinreichendes Kriterium für Konfluenz

Definition 6.1.4 ($\twoheadrightarrow\!\!\!\!\!\shortmid\;$-Substitutionsstabilität) Sei ein HRS R gegeben und die Relationen $\rightarrow$ und $\twoheadrightarrow\!\!\!\!\!\shortmid\;$ durch R induziert. Ein λ-Term t in langer $\beta\eta$-Normalform heiße **$\twoheadrightarrow\!\!\!\!\!\shortmid\;$-substitutionsstabil** genau dann, wenn für beliebige Substitutionen σ, σ' gilt

$$\sigma \twoheadrightarrow\!\!\!\!\!\shortmid\;\; \sigma' \;\Rightarrow\; \sigma(t)\downarrow_\beta \twoheadrightarrow\!\!\!\!\!\shortmid\;\; \sigma'(t)\downarrow_\beta$$

wobei $\sigma \twoheadrightarrow\!\!\!\!\!\shortmid\;\; \sigma'$ wie üblich bedeutet

$$Dom(\sigma) \subseteq Dom(\sigma') \;\wedge\; \forall F \in Dom(\sigma).\ \sigma(F) \twoheadrightarrow\!\!\!\!\!\shortmid\;\; \sigma'(F)$$

Ein HRS heiße $\twoheadrightarrow\!\!\!\!\!\shortmid\;$-substitutionsstabil, wenn alle rechten Seiten seiner Regeln es sind.

Bemerkungen:

- Diese Bedingung ist für Termersetzungssysteme erster Ordnung stets erfüllt da alle Variablen von Basistyp sind. Für Termersetzungssysteme höherer Ordnung stellt sie jedoch eine echte Einschränkung dar. Wie später an einem Beispiel gezeigt werden wird gilt das folgende Lemma 6.1.6 ohne diese Einschränkung nicht.

- Da in Termersetzungsregeln die linken Seiten Patterns sein müssen ist die Bedingung für die linken Seiten stets erfüllt.

Beispiel 6.1.5 Nicht jeder Term höherer Ordnung ist $\twoheadrightarrow\!\!\!\!\!\!\mid\,\,\,$-substitutionsstabil. Sei R das HRS mit der einzigen Regel

$$g(G) \rightarrow h(G)$$

Seien F und S Variablen mit den Typen

$$\begin{array}{lcl} F & : & (\tau_1 \rightarrow \tau_2) \rightarrow \tau_2 \\ S & : & \tau_1 \rightarrow \tau_2 \end{array}$$

wobei τ_1 und τ_2 Basistypen sind.

Seien außerdem zwei Substitutionen θ und θ' gegeben mit

$$\begin{array}{lcl} \theta(F) & = & \lambda x.x(x(c)) \\ \theta(S) & = & \lambda x.g(x) \\ \theta'(F) & = & \lambda x.x(x(c)) \\ \theta'(S) & = & \lambda x.h(x) \end{array}$$

Dann gilt zwar $\theta \twoheadrightarrow\!\!\!\!\!\!\mid\,\,\, \theta'$ aber **nicht**

$$\theta(F(S))\downarrow_\beta = g(g(c)) \twoheadrightarrow\!\!\!\!\!\!\mid\,\,\, h(h(c)) = \theta'(F(S))\downarrow_\beta$$

Lemma 6.1.6 *Sei R ein HRS, das parallel abgeschlossen, linkslinear und $\twoheadrightarrow\!\!\!\!\!\!\mid\,\,\,$-substitutionsstabil ist. Dann hat die von R induzierte parallele Reduktion $\twoheadrightarrow\!\!\!\!\!\!\mid\,\,\,_R$ die Diamanteigenschaft.*

Beweis Es gelte $M \twoheadrightarrow\!\!\!\!\!\!\mid\,\,\,^{\mathcal{U}} N_1$ und $M \twoheadrightarrow\!\!\!\!\!\!\mid\,\,\,^{\mathcal{V}} N_2$. Dabei sind $\mathcal{U}$ und $\mathcal{V}$ die verwendeten Redexpositionen in den parallelen Reduktionen von M nach N_1 bzw. N_2. $\mathcal{U}$ und $\mathcal{V}$ sind also endliche, i.a. nicht gleichmächtige Teilmengen von $Pos(M)$. Damit werden als von $\mathcal{U}$ und $\mathcal{V}$ abhängige Größen definiert:

$$\begin{array}{lcl} \mathcal{W} & := & \{u \in \mathcal{U} \mid \exists v \in \mathcal{V}.v \leq u\} \cup \{v \in \mathcal{V} \mid \exists u \in \mathcal{U}.u \leq v\} \\ \overline{\mathcal{W}} & := & ((\mathcal{U} \cup \mathcal{V}) - \mathcal{W}) \cup (\mathcal{U} \cap \mathcal{V}) \end{array}$$

Gezeigt wird nun

$$\exists N.\ N_1 \twoheadrightarrow\!\!\!\!\!\!\mid\,\,\, N \twoheadleftarrow\!\!\!\!\!\!\mid\,\,\, N_2$$

durch vollständige Induktion über

$$\mathcal{L}(M,\mathcal{U},\mathcal{V}) := \sum_{w \in \mathcal{W}} L(M/w)$$

Induktionsanfang: Sei $\mathcal{L} = 0$.
Wegen $\forall t \in \mathcal{T}.L(t) > 0$ folgt $W = \emptyset$, und daher

$$(*) \quad \forall u \in \mathcal{U}\ \forall v \in \mathcal{V}.\ u \parallel v$$

Seien $(l_{u_i} \to r_{u_i})$ bzw. $(l_{v_i} \to r_{v_i})$ die an den Positionen $u_i \in \mathcal{U}$ bzw. $v_i \in \mathcal{V}$ angewendeten Regeln und θ_{u_i} bzw. θ_{v_i} die verwendeten Substitutionen. Dann gilt

$$\begin{array}{rcl} N_1 & = & N[u_i \leftarrow \theta_{u_i}(r_{u_i})\downarrow_\beta \mid u_i \in \mathcal{U}] \\ N_2 & = & N[v_i \leftarrow \theta_{v_i}(r_{v_i})\downarrow_\beta \mid v_i \in \mathcal{V}] \end{array}$$

Aus der Eigenschaft $(*)$ folgt

$$N_1[v_i \leftarrow \theta_{v_i}(r_{v_i})\downarrow_\beta \mid v_i \in \mathcal{V}] \;=\; N_2[u_i \leftarrow \theta_{u_i}(r_{u_i})\downarrow_\beta \mid u_i \in \mathcal{U}] \;=:\; N$$

und daher

$$N_1 \overset{\mathcal{V}}{\twoheadrightarrow} N \overset{\mathcal{U}}{\twoheadleftarrow} N_2$$

Induktionsschluß: Da es maximale Elemente in $\mathcal{U} \cup \mathcal{V}$ gibt ist $\overline{\mathcal{W}} \neq \emptyset$. Wähle ein beliebiges $u \in \overline{\mathcal{W}}$, o.B.d.A. $u \in \mathcal{U}$. Es wird zunächst gezeigt:

$$\exists M_u.\ N_1/u \twoheadrightarrow M_u \twoheadleftarrow N_2/u$$

Sei $V_u := \{v \in \mathcal{V} \mid u \leq v\}$ und $(l \to r) \in R$ die an der Position u mit Substitution σ angewendete Regel. Da nach Voraussetzung alle Positionen in $\mathcal{U}$ parallel sind folgt $M/u = \sigma(l)\downarrow_\beta$ und $N_1/u = \sigma(r)\downarrow_\beta$. Es sind zwei Fälle zu unterscheiden:

1. Es gelte

$$(**) \quad \forall v \in V_u.\ v/u = w \cdot w'$$

   wobei l/w von der Form $F(\overline{x_k})$ ist. Es wird definiert

$$\{F_1, \ldots, F_n\} := \{F \in fv(l) \mid \exists v \in V_u.\ v/u = w \cdot w' \wedge l/w = F(\ldots)\}$$

   Da nach Voraussetzung alle Regeln in R linkslinear sind kommt jedes F_i nur einmal in l vor. Deswegen gehört zu jedem F_i eindeutig ein w_i mit $l/w_i = F_i(\ldots)$. Sei W_i' die Menge aller zu F_i gehörenden w', d.h.

$$W_i' := \{w_{i,1}', \ldots, w_{i,n_i}'\} := \{v/(u \cdot w_i) \mid \exists v \in V_u.\ u \cdot w_i \leq v\}$$

   Da nach Voraussetzung alle Elemente von $\mathcal{V}$ parallel sind gilt dies auch für alle Elemente von W_i'. Da l ein Pattern ist gilt $\forall i, j.\ w_i \parallel w_j$. Daher ist die folgende Vereinigung disjunkt und es folgt mit $(**)$

$$\bigcup_{i=1}^{n} u \cdot w_i \cdot W_i' \;=\; V_u$$

   Da nach Konvention alle Elemente von $Cod(\sigma)$ in langer $\beta\eta$-Normalform sein müssen ist $\sigma(F_i) = \lambda\overline{x_{k_i}}.t_i$. Es seien für $i = 1, \ldots, n$ und $k = 1, \ldots, n_i$

die an den Positionen $u{\cdot}w_i{\cdot}w'_{i,k} \in V_u$ mit den Substitutionen $\sigma_{i,k}$ angewendeten Regeln mit $(l_{i,k} \to r_{i,k})$ bezeichnet. Damit wird eine neue Substitution definiert:

$$\sigma' := \sigma + \sum_{i=1}^{n} \{F_i \mapsto \lambda\overline{x_{k_i}}.(t_i[w'_{i,k} \leftarrow \sigma_{i,k}(r_{i,k})\downarrow_\beta \mid k = 1, \ldots, n_i]\}$$

Es gilt:

$$\begin{array}{rcl} t_i/w'_{i,k} & = & \sigma(F_i(\overline{x_{k_i}}))\downarrow_\beta/w'_{i,k} \\ & = & \sigma(l/w_i)\downarrow_\beta/w'_{i,k} \\ & = & (\sigma(l)\downarrow_\beta/w_i)/w'_{i,k} \\ & = & \sigma(l)\downarrow_\beta/(w_i \cdot w'_{i,k}) \\ & = & \sigma_{i,k}(l_{i,k})\downarrow_\beta \end{array}$$

und daher $\sigma \mathrel{-\!\!\shortparallel\!\!\rightarrow} \sigma'$. Da r nach Voraussetzung $\mathrel{-\!\!\shortparallel\!\!\rightarrow}$-substitutionsstabil ist folgt nach Def. 6.1.4

$$N_1/u \;=\; \sigma(r)\downarrow_\beta \;\mathrel{-\!\!\shortparallel\!\!\rightarrow}\; \sigma'(r)\downarrow_\beta$$

Da R nach Voraussetzung linkslinear ist kommt jede der Variablen F_i nur *einmal* in l vor, woraus folgt

$$N_2/u \;=\; \sigma'(l)\downarrow_\beta \;\rightarrow\; \sigma'(r)\downarrow_\beta$$

Definiere nun $M_u := \sigma'(r)\downarrow_\beta$, dann gilt

$$N_1/u \mathrel{-\!\!\shortparallel\!\!\rightarrow} M_u \mathrel{\leftarrow\!\!\shortparallel\!\!-} N_2/u$$

und somit ist die Behauptung gezeigt.

2. Wenn die Bedingung des ersten Falles nicht erfüllt war, dann gilt

$$\exists v_1 \in V_u.\; w := v_1/u \in Pos(l) \;\wedge\; l/w \neq F(\ldots)$$

Sei $(l_1 \to r_1)$ die an v_1 angewendete Regel und σ_1 die verwendete Substitution. Es gilt $\sigma(l/w)\downarrow_\beta = M/v_1 = \sigma_1(l_1)\downarrow_\beta$. Da o.B.d.A. $bv(l) \cap Cod(\sigma) = \emptyset$ ist gibt es nach Lemma 2.2.3 ein kritisches Paar $\langle Q_1, Q_2\rangle$ und eine Substitution δ mit

$$\begin{array}{rcccl} N_1/u & = & \sigma(r)\downarrow_\beta & = & \delta(Q_1)\downarrow_\beta \\ M/u[\sigma_1(r_1)\downarrow_\beta]_w & = & \sigma(l)\downarrow_\beta[\sigma_1(r_1)\downarrow_\beta]_w & = & \delta(Q_2)\downarrow_\beta \end{array}$$

Sei $\hat{M} := \delta(Q_2)\downarrow_\beta$. Weil R nach Voraussetzung parallel abgeschlossen ist folgt

$$\hat{M} \mathrel{-\!\!\shortparallel\!\!\rightarrow}^{\hat{W}} N_1/u$$

wobei $\hat{W}$ die Menge der Redexpositionen in der parallelen Reduktion von $\hat{M}$ nach N_1/u sei. Im Übergang von M/u nach $\hat{M}$ wurde einer der durch die Positionen in V_u bestimmten Redexe reduziert. Wenn alle diese Redexe reduziert werden, so gelangt man von M/u nach N_2/u. Da alle Positionen in $\mathcal{V}$ und daher auch in V_u parallel sind beeinflussen diese Reduktionen sich gegenseitig nicht. Sei also V' die Menge der restlichen Redexpositionen in $\hat{M}$, d.h.

$$V' := \{v/u \mid v \in V_u - \{v_1\}\}$$

Dann folgt daraus

$$N_1/u \xleftarrow{\hat{W}} \hat{M} \xrightarrow{V'} N_2/u$$

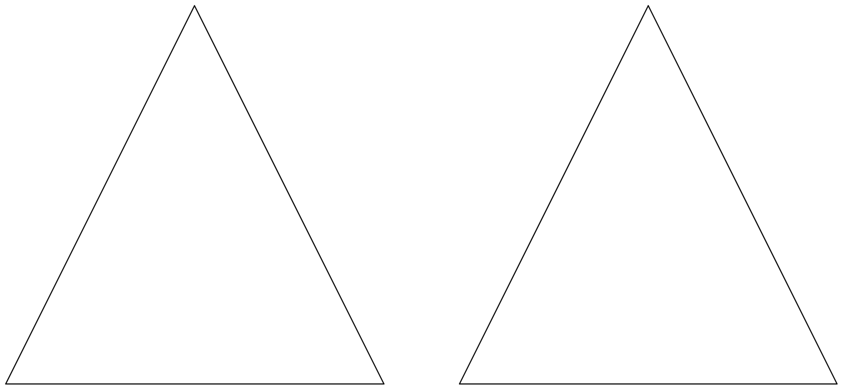
Es wird nun gezeigt, daß $\mathcal{L}(\hat{M}, V', \hat{W}) < \mathcal{L}(M, \mathcal{U}, \mathcal{V})$ ist, wodurch die Induktionshypothese anwendbar wird.

Sei der Wert $\mathcal{L}'$ definiert als

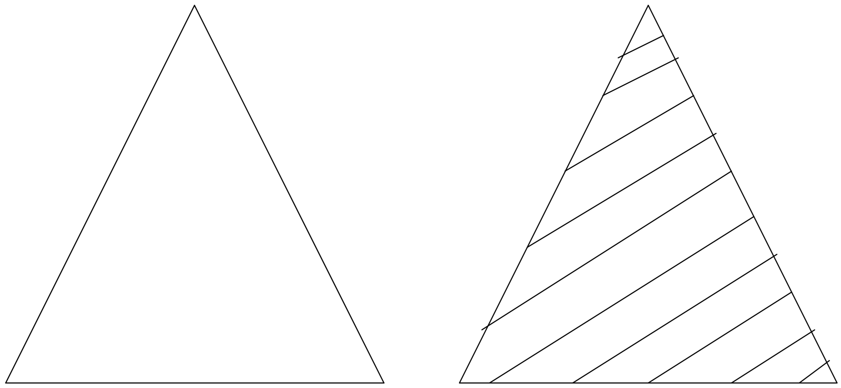
$$\mathcal{L}' := \sum_{v' \in V'} L(M/(u \cdot v')) = \sum_{v \in V_u - \{v_1\}} L(M/v)$$

Es wird zunächst $\mathcal{L}(\hat{M}, V', \hat{W}) \leq \mathcal{L}'$ gezeigt. Dabei sind mehrere Fälle zu unterscheiden. Sei im folgenden $v', v'_1, \ldots, v'_k \in V'$ und $\hat{w}, \hat{w}_1, \ldots, \hat{w}_n \in \hat{W}$. Es ist wichtig, zu beachten, daß sowohl alle Elemente von V' als auch alle Elemente von $\hat{W}$ zueinander parallel sind. In den folgenden vier Abbildungen wird auf der linken Seite der Beitrag zu $\mathcal{L}(\hat{M}, V', \hat{W})$ und auf der rechten Seite der Beitrag zu $\mathcal{L}'$ schraffiert dargestellt.

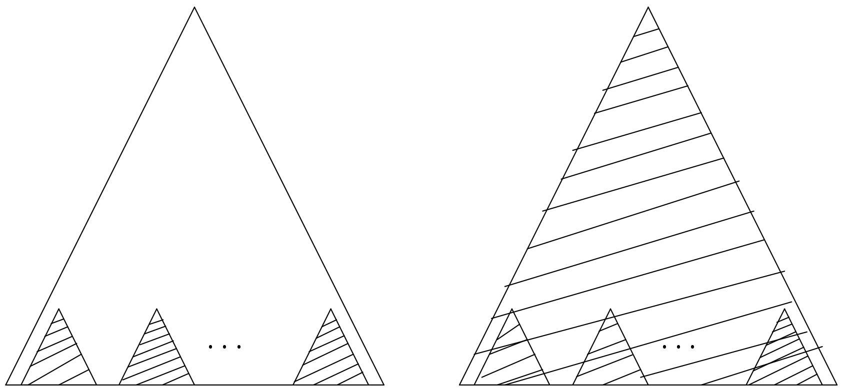
(a) $\hat{w} \parallel V'$
In diesem Fall zählt die Größe $L(\hat{M}/\hat{w})$ des Subterms $\hat{M}/\hat{w}$ weder zu $\mathcal{L}(\hat{M}, V', \hat{W})$ noch zu $\mathcal{L}'$.



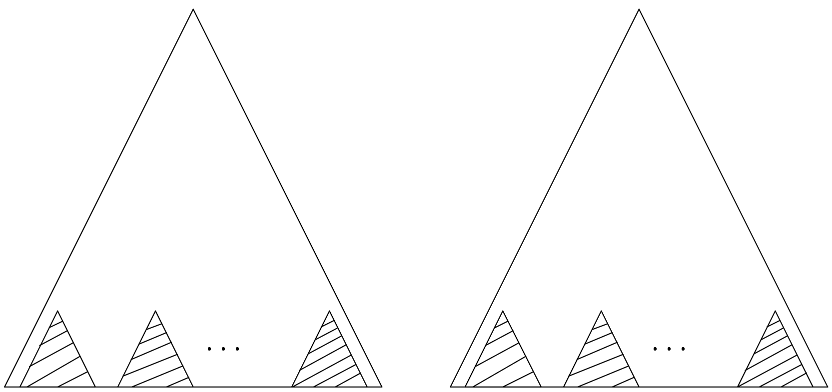
(b) $v' \parallel \hat{W}$
In diesem Fall zählt die Größe $L(\hat{M}/v')$ des Subterms $\hat{M}/v'$ mit zu $\mathcal{L}'$ aber nicht zu $\mathcal{L}(\hat{M}, V', \hat{W})$.



(c) $v' \leq \hat{w}_1, \ldots, \hat{w}_k$ für ein $k \in \mathbb{N}$.
Hier zählt $\sum_{i=1}^{k} L(\hat{M}/\hat{w}_i)$ zu $\mathcal{L}(\hat{M}, V', \hat{W})$ und $L(\hat{M}/v')$ zu $\mathcal{L}'$. Nach der Bemerkung zu Def. 1.1.1 gilt, daß letzteres größer (oder gleich) als ersteres ist, und somit trägt der Subterm $\hat{M}/v'$ gleichviel oder mehr zu $\mathcal{L}'$ bei als zu $\mathcal{L}(\hat{M}, V', \hat{W})$.



(d) $\hat{w} \leq {v_1}', \ldots, {v_k}'$ für ein $k \in \mathbb{N}$.
Hier trägt der Subterm $\hat{M}/\hat{w}$ den gleichen Wert zu $\mathcal{L}(\hat{M}, V', \hat{W})$ und zu $\mathcal{L}'$ bei, nämlich $\sum_{i=1}^{k} L(\hat{M}/{v_i}')$.



Somit gilt, daß jeder Subterm von $\hat{M}$ gleichviel oder mehr zu $\mathcal{L}'$ beiträgt als zu $\mathcal{L}(\hat{M}, V', \hat{W})$. Daraus folgt die Behauptung $\mathcal{L}(\hat{M}, V', \hat{W}) \leq \mathcal{L}'$. Da nach der Bemerkung zu Def. 1.1.1 $L(\hat{M}/v_1) > 0$ ist folgt

$$\mathcal{L}' := \sum_{v \in V_u - \{v_1\}} L(M/v) < \sum_{v \in V_u} L(M/v)$$

Da $V_u \subseteq W$ ist gilt außerdem

$$\sum_{v \in V_u} L(M/v) \leq \mathcal{L}(M, \mathcal{U}, \mathcal{V})$$

und somit

$$\mathcal{L}(\hat{M}, V', \hat{W}) \leq \mathcal{L}' < \sum_{v \in V_u} L(M/v) \leq \mathcal{L}(M, \mathcal{U}, \mathcal{V})$$

Daher ist die Induktionshypothese anwendbar, woraus folgt, daß es ein M_u gibt mit $N_2/u \twoheadrightarrow\!\!\!\!\shortmid\;\; M_u \twoheadleftarrow\!\!\!\!\shortmid\;\; N_1/u$. Sei nun

$$N := M[u \leftarrow M_u \;\mid\; u \in \overline{\mathcal{W}}]$$

Da alle Elemente von $\overline{\mathcal{W}}$ zueinander parallel sind, und da nach Def. von $\overline{\mathcal{W}}$ gilt $\forall p \in (\mathcal{U} \cup \mathcal{V})\ \exists q \in \overline{\mathcal{W}}.\ q \leq p$ folgt

$$N_1 \;\twoheadrightarrow\!\!\!\!\shortmid\;\; N \;\twoheadleftarrow\!\!\!\!\shortmid\;\; N_2$$

Somit hat $\twoheadrightarrow\!\!\!\!\shortmid\;$ die Diamanteigenschaft.

□

Die Voraussetzung der $\twoheadrightarrow\!\!\!\!\shortmid\;$-Substitutionsstabilität ist für das gerade bewiesene Lemma unbedingt nötig, wie das folgende Beispiel zeigt.

Beispiel 6.1.7 Sei R das HRS mit den folgenden Regeln:

1. $f(\lambda y.F(y)) \rightarrow F(\lambda x.g(x))$
2. $k(\lambda y.F(y)) \rightarrow F(\lambda x.h(x))$
3. $g(F) \rightarrow l(F)$
4. $h(F) \rightarrow l(F)$

Der Term T wird definiert als

$$T := e(f(\lambda x.x(x(c))), k(\lambda x.x(x(c))))$$

Daraus folgt

$$T \twoheadrightarrow\!\!\!\!\shortmid\;\; e(g(g(c)), k(\lambda x.x(x(c)))) := T_1$$

und

$$T \mathrel{-\!\!\shortparallel\!\!\twoheadrightarrow} e(f(\lambda x.x(x(c))), h(h(c))) := T_2$$

T_1 und T_2 haben zwar das gemeinsame Redukt

$$e(l(l(c)), l(l(c))$$

doch kann dieses weder von T_1 noch von T_2 aus in einem $\mathrel{-\!\!\shortparallel\!\!\twoheadrightarrow}$-Schritt erreicht werden. Somit hat $\mathrel{-\!\!\shortparallel\!\!\twoheadrightarrow}$ für dieses HRS nicht die Diamanteigenschaft, obwohl es keine kritischen Paare hat. Im nächsten Abschnitt über die Konfluenz semiorthogonaler HRS wird jedoch ein Satz bewiesen, aus dem folgt, daß dieses HRS trotzdem konfluent ist.

Satz 6.1.8 *Sei R ein HRS, das $\mathrel{-\!\!\shortparallel\!\!\twoheadrightarrow}$-substitutionsstabil, parallel abgeschlossen und linkslinear ist. Dann ist die von R induzierte Relation $\rightarrow_R$ konfluent.*
Beweis direkt aus Lemma 6.1.6, der Bemerkung zu Def. 6.1.1 und Lemma 2.1.8.
□

Beispiel 6.1.9 Ein Beispiel für die Anwendung dieses Satzes ist das folgende HRS. Seine Regeln beschreiben die η-Konversion in einem erweiterten λ-Kalkül, in dem zusätzlich das Symbol $\perp$ für den total undefinierten λ-Term eingeführt wurde. Dieses HRS R besteht aus folgenden drei Regeln:

1. $abs(\lambda x.app(T, x)) \rightarrow T$
2. $app(\perp, T) \rightarrow \perp$
3. $abs(\lambda x.\perp) \rightarrow \perp$

Dabei können nur die erste und die zweite Regel miteinander überlappen, was zu dem kritischen Paar $\langle \perp, abs(\lambda x.\perp)\rangle$ führt. Da in diesem kritischen Paar keine freien Variablen vorkommen folgt aus der dritten Regel sofort die parallele Abgeschlossenheit von R. Eine freie Variable in der rechten Seite einer Regel kommt nur bei der ersten Regel in der einfachen Form T vor. Daher ist die Voraussetzung der $\mathrel{-\!\!\shortparallel\!\!\twoheadrightarrow}$-Substitutionsstabilität trivialerweise erfüllt. Da R außerdem linkslinear ist folgt die Konfluenz von $\rightarrow_R$ aus dem Satz 6.1.8.

6.2 Semiorthogonale Systeme

Das im letzten Abschnitt gezeigte Ergebnis ist zwar für Termersetzungssysteme erster Ordnung sehr brauchbar, die Verallgemeinerung auf HRS ist jedoch nur beschränkt anwendbar, da die Eigenschaft der $\mathrel{-\!\!\shortparallel\!\!\twoheadrightarrow}$-Substitutionsstabilität hier häufig nicht erfüllt ist.

Daher wird versucht, Ergebnisse zu finden, die ohne diese Voraussetzung auskommen. Dabei werden jedoch stärkere Forderungen an die kritischen Paare eines HRS gestellt als die parallele Abgeschlossenheit. Statt der im letzten Abschnitt definierten parallelen Reduktion wird hier die im zweiten Kapitel definierte parallele bottom-up Reduktion verwendet werden.

6.2.1 Orthogonalität

Definition 6.2.1 (Orthogonal) Ein HRS R heißt **orthogonal**, wenn es linkslinear ist und keine zwei Regeln aus R überlappen, d.h. wenn R keine kritischen Paare hat.

Definition 6.2.2 (Semiorthogonal) Sei R ein linkslineares HRS mit den Regeln $(l \rightarrow r), (l_1 \rightarrow r_1), \ldots, (l_n \rightarrow r_n)$ (die *nicht* paarweise verschieden sein müssen). Es gebe außerdem Positionen $p_1, \ldots, p_n \in Pos(l)$ und Substitutionen $\theta, \theta_1, \ldots, \theta_n$ für die gilt

- $p_1 < p_2 < \ldots < p_n$
- $\forall 1 \leq i \leq n.\ \theta(l)\downarrow_\beta / p_i = \theta_i(l_i)\downarrow_\beta$
- Für alle $i \in \{1, \ldots, n\}$ überlappt die Regelanwendung $(l_i \rightarrow r_i)$ an p_i mit der Regelanwendung $(l \rightarrow r)$ an ϵ.
- $\forall 1 \leq i \leq n-1$ überlappt die Regelanwendung $(l_{i+1} \rightarrow r_{i+1})$ an p_{i+1} nicht mit der Regelanwendung $(l_i \rightarrow r_i)$ an p_i.

Man nennt dann $(\theta_1(l_1)\downarrow_\beta, \ldots, \theta_n(l_n)\downarrow_\beta)$ einen **kritischen Turm** der Höhe n in $\theta(l)\downarrow_\beta$ (siehe Abbildung 6.2). Da nach Voraussetzung für $j > i$ die Regelanwendung an p_j unkritisch zu der an p_i ist beeinflußt die Anwendung der Regel $(l_j \rightarrow r_j)$ an p_j die Anwendbarkeit der Regel $(l_i \rightarrow r_i)$ an p_i nicht.

Es wird nun gefordert, daß der Term, der aus $\theta(l)\downarrow_\beta$ entsteht, wenn man nacheinander von unten nach oben die jeweiligen Termersetzungsregeln an den Positionen $p_n, p_{n-1}, \ldots, p_1$ anwendet, gleich dem Term $\theta(r)\downarrow_\beta$ ist.

Ein HRS heißt **semiorthogonal**, wenn es diese Bedingungen erfüllt.

Mit Hilfe der im zweiten Kapitel in Def. 2.2.6 definierten parallelen bottom-up Reduktion und Lemma 2.2.9 läßt sich diese Eigenschaft noch einfacher darstellen. Ein HRS ist semiorthogonal genau dann, wenn es linkslinear ist und folgende Eigenschaft gilt:

Sei $(l \rightarrow r) \in R$, θ eine Substitution und $\mathcal{M} \subseteq Pos(l)$ eine *total* geordnete Menge von Positionen von kritischen Redexen in $\theta(l)\downarrow_\beta$. Dann gilt

$$\theta(l)\downarrow_\beta >^{\mathcal{M}} t \quad \Rightarrow \quad t = \theta(r)\downarrow_\beta$$

Definition 6.2.3 (Schwache Orthogonalität) Ein HRS heißt **schwach orthogonal**, wenn es linkslinear ist und für alle seine kritischen Paare $\langle u_1, u_2 \rangle$ gilt: $u_1 = u_2$.

Lemma 6.2.4 *Aus diesen Definitionen und der Definition der kritischen Paare in Def. 2.2.2 folgt direkt:*

$$\textit{orthogonal} \Rightarrow \textit{semiorthogonal} \Rightarrow \textit{schwach orthogonal}$$

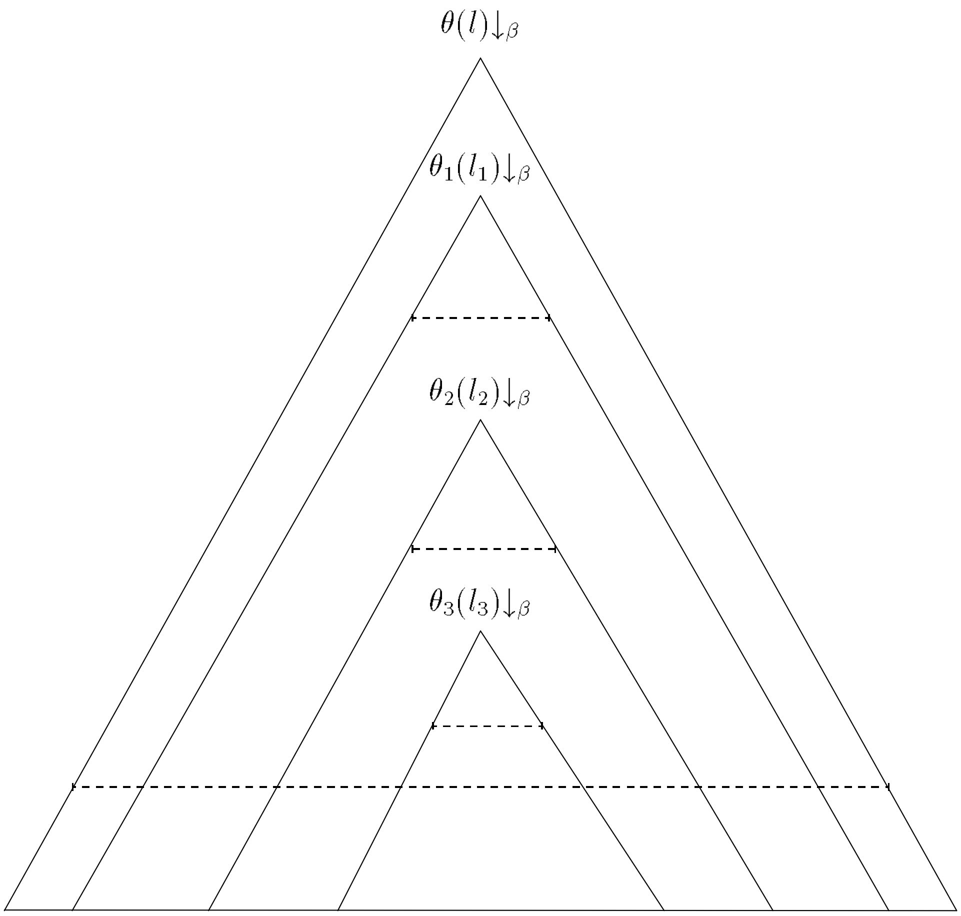


Abbildung 6.2: Ein kritischer Turm der Höhe 3 in $\theta(l)\downarrow_\beta$. Dabei ist der Teil des Redexes über der gestrichelten Linie durch die Termersetzungsregel festgelegt. Der Teil darunter entspricht den durch die Substitution in die freien Variablen der linken Seite der Regel eingesetzten Termen. Ein Redex ist somit kritisch, bzw. unkritisch zu dem darüberliegenden Redex wenn er über, bzw. unter der gestrichelten Linie des darüberliegenden Redexes beginnt.

Die Konfluenz orthogonaler HRS wurde von Nipkow in [3] bewiesen. Ein noch weitergehendes Ergebnis über die Konfluenz semiorthogonaler HRS, das von Takahashi [6] bewiesen wurde, wird im nächsten Abschnitt dargestellt. Das folgende technische Lemma wird in diesem Beweis benötigt.

Lemma 6.2.5 *Sei R ein schwach orthogonales HRS mit Regeln $(l \to r), (l_1 \to r_1)$ und $(l_2 \to r_2)$, und seien t, t_1, t_2 drei Redexe mit $t = \theta(l)\downarrow_\beta$, $t_1 = \theta_1(l_1)\downarrow_\beta = t/p_1$ und $t_2 = \theta_2(l_2)\downarrow_\beta = t/p_2$. Sei außerdem $p_1 \parallel p_2$ und sowohl t_1 als auch t_2 kritisch zu t. (Abbildung 6.3 zeigt diesen Fall). Dann sind alle drei Redexe gleich ihrem Kontraktum.*

Beweis Da t_1 und t_2 kritisch zu t sind und R schwach orthogonal ist gilt

$$\theta(l)\downarrow_\beta[\theta_1(r_1)\downarrow_\beta]_{p_1} = \theta(r)\downarrow_\beta = \theta(l)\downarrow_\beta[\theta_2(r_2)\downarrow_\beta]_{p_2}$$

Da $p_1 \parallel p_2$ ist folgt daraus $\theta_1(r_1)\downarrow_\beta = t/p_1$ und $\theta_2(r_2)\downarrow_\beta = t/p_2$ und daher auch $\theta(r)\downarrow_\beta = t$. □

Bemerkung: Da aus der Semiorthogonalität die schwache Orthogonalität folgt gilt dieses Lemma erst recht für semiorthogonale HRS.

6.2.2 Semiorthogonale HRS sind konfluent

Der folgende Satz über die Konfluenz semiorthogonaler HRS ist im wesentlichen eine Übertragung eines Beweises von Takahashi [6] über die Konfluenz semiorthogonaler CLC (CLC = conditional λ-calculi). Außer durch andere notationelle Konventionen unterscheiden sich CLC von HRS hauptsächlich dadurch, daß zu den Termersetzungsregeln auch noch Bedingungen angegeben werden, die erfüllt sein müssen, damit eine Reduktion stattfinden kann. Diese Erweiterung ist jedoch insofern entbehrlich, als für jedes Regelsystem mit Bedingungen ein anderes ohne Bedingungen gefunden werden kann, das das erste simuliert.

Außerdem wurden im Beweis von Takahashi andere Hilfssätze verwendet als in den Arbeiten von Nipkow, und es wurde eine zusätzliche Voraussetzung verwendet, die jedoch für den Beweis der Konfluenz unnötig ist.

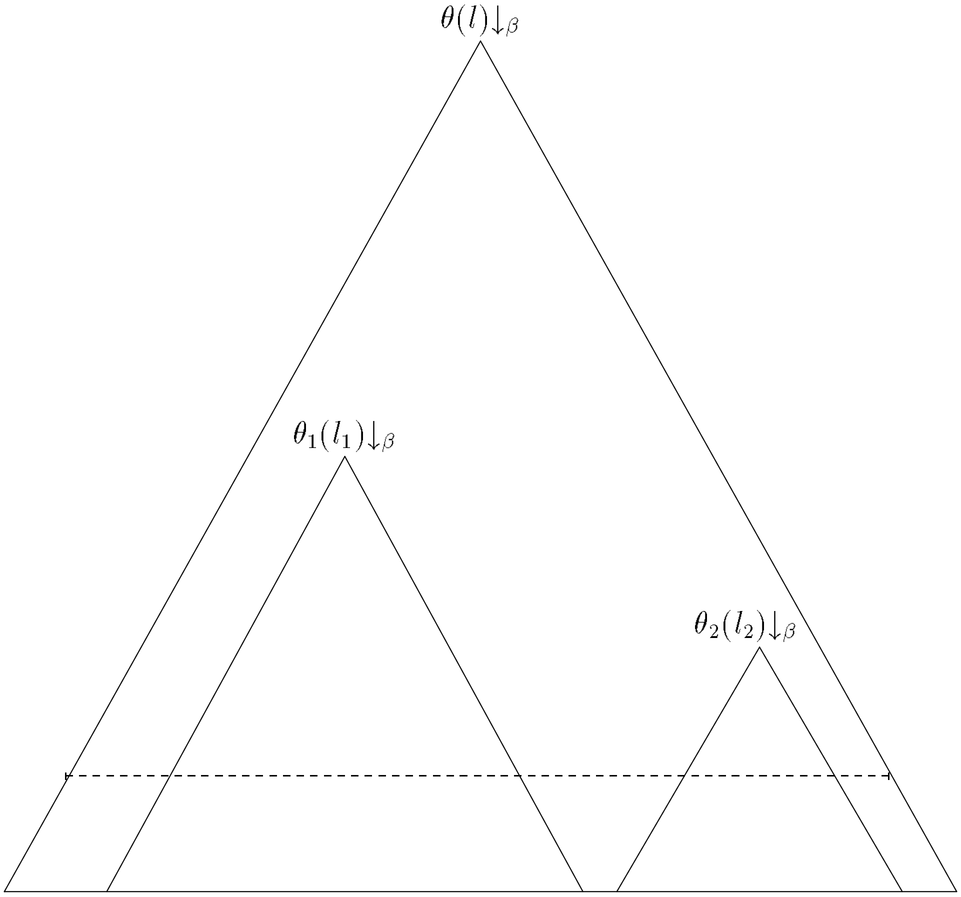
Hier wird dieser Beweis noch einmal mit Hilfe der im zweiten Kapitel bewiesenen Lemmata in der Notation von Nipkow ohne diese unnötige Voraussetzung dargestellt.

Definition 6.2.6 Sei ein HRS R gegeben. Sei n die Anzahl der Termersetzungsregeln in R. Numeriere nun die Regeln in R willkürlich von 1 bis n, d.h. wähle eine beliebige Bijektion

$$Num : \begin{cases} R & \mapsto \quad \{1, \ldots, n\} \\ (l \to r) \in R & \mapsto \quad Num(l \to r) \end{cases}$$

Für jeden λ-Term t in langer $\beta\eta$-Normalform wird der λ-Term t^* rekursiv wie folgt definiert:

Abbildung 6.3: Parallele kritische Redexe



1. $a^* = a$ für alle Atome a, falls a kein Redex ist.

2. $(\lambda x.t)^* = \lambda x.(t^*)$

3. $(A\ B)^* = (A^*\ B^*)$ falls $(A\ B)$ kein Redex ist

4. Falls es Regeln $(l \to r) \in R$ und Substitutionen θ gibt für die $\theta(l){\downarrow_\beta} = t$ ist, so wähle unter diesen diejenige mit dem kleinsten Index aus. Sei diese Regel $(l' \to r')$. Formal wird dies definiert durch:

$$\begin{array}{l} K := \{(l \to r) \in R \mid \exists \theta.\ \theta(l){\downarrow_\beta} = t\} \\ (l' \to r') \in K \\ \forall (l \to r) \in K.\ Num(l' \to r') \leq Num(l \to r) \end{array}$$

   Diese Definition ist eindeutig, weil Num eine Bijektion ist. Dann wird t^* definiert als
$$t^* = \theta^*(r'){\downarrow_\beta}$$
   wobei
$$\forall F \in fv(l).\ \theta^*(F) = (\theta(F))^*$$
   und θ^* auf $\mathcal{T}$ fortgesetzt wird wie im ersten Kapitel definiert.

Es bleibt zu zeigen, daß der Term t^* wohldefiniert ist. Weil die linke Seite l einer Regel keine freie Variable sein darf ist $\theta(F)$ ein *echter* Subterm von $\theta(l){\downarrow_\beta}$ und somit schon rekursiv definiert.

Im Gegensatz zu dieser Definition wurde bei Takahashi [6] im vierten Fall nur die Definition
$$(\theta(l){\downarrow_\beta})^* = \theta^*(r){\downarrow_\beta}$$
gemacht. Da es vorkommen kann, daß zwei verschiedene Regeln aus R an Position ϵ überlappen, ist t^* so i.a. nicht wohldefiniert. Bei Takahashi wurde daher zusätzlich vorausgesetzt, daß zwei verschiedene Regeln aus R nicht an Position ϵ überlappen dürfen, wodurch t^* wieder wohldefiniert wird. Diese zusätzliche Voraussetzung ist z.B. für orthogonale HRS immer erfüllt. Zwar folgt diese Voraussetzung nicht aus der Semiorthogonalität, jedoch ist das so definierte t^* für semiorthogonale HRS möglicherweise trotzdem wohldefiniert. Es könnte sein, daß aus der Semiorthogonalität folgt, daß obwohl manchmal mehrere verschiedene Regeln zur Auswahl stehen, das Endergebnis immer das gleiche ist. Das würde bedeuten, daß in der ersten Definition von t^* die gewählte Numerierung Num der Regeln in R keinen Einfluß auf das Ergebnis hat.

Im folgenden wird jedoch die erste Definition von t^* verwendet, die auch ohne zusätzliche Voraussetzungen sicher wohldefiniert ist.

Lemma 6.2.7 *Sei R ein semiorthogonales HRS, Num eine Numerierung der Regeln in R, $>$ die von R induzierte parallele bottom-up Reduktion und t ein λ-Term in langer $\beta\eta$-Normalform. Dann gilt*

$$t > t' \quad \Rightarrow \quad t' > t^*$$

wobei t^* wie in Def. 6.2.6 mit Hilfe von Num definiert ist.
Beweis durch Induktion über den Aufbau von t.

1. Sei t *kein* Redex, d.h. $\nexists (l \to r) \in R, \theta.\ t = \theta(l)\downarrow_\beta$

   (a) $t = a$, a ein Atom $\Rightarrow \quad t' = a = t^*$

   (b) $t = \lambda x.d$. Daraus folgt $t' = \lambda x.d'$ mit $d > d'$. Aus der Induktionshypothese folgt $d' > d^*$ und somit $t' = \lambda x.d' > \lambda x.d^* = t^*$.

   (c) $t = (A\ B)$. Daraus folgt $t' = (A'\ B')$ mit $A > A'$ und $B > B'$. Aus der Induktionshypothese folgt $A' > A^*$ und $B' > B^*$ und somit $t' = (A'\ B') > (A^*\ B^*) = t^*$.

2. Jetzt sei t ein Redex, d.h.

   $$K := \{(l \to r) \in R \ \mid\ \exists \theta.\ \theta(l)\downarrow_\beta = t\} \neq \emptyset$$

   Sei $(l' \to r') \in K$ die Regel aus K mit dem kleinsten Index bezüglich der Numerierung Num. Es treten wieder mehrere Fälle auf.

   (a) $t = \theta(l')\downarrow_\beta > \theta'(r')\downarrow_\beta = t'$ mit $\theta > \theta'$. Aus der Induktionshypothese folgt $\theta' > \theta^*$ und somit

   $$t' = \theta'(r')\downarrow_\beta > \theta^*(r')\downarrow_\beta = t^*$$

   nach Lemma 2.2.7, Lemma 2.2.11 und der Definition von t^*.

   (b) $t = a$, a ein Atom. Daraus folgt $t' > \theta(r)\downarrow_\beta = t^*$ oder $t' = \theta(r)\downarrow_\beta = t^*$. Die Eindeutigkeit folgt aus der Semiorthogonalität.

   (c) $t > t'$, wobei nicht die Regel $(l' \to r')$ an Position $\epsilon \in Pos(t)$ angewandt wurde. Sei $\mathcal{U}$ die Menge der im Übergang $t > t'$ kontrahierten Redexe. Sei $\mathcal{U}' \subseteq \mathcal{U}$ die Teilmenge der Redexe die zu der Regelanwendung von $(l' \to r')$ an der Position $\epsilon \in Pos(t)$ kritisch sind. Darunter kann auch eine Anwendung einer Regel aus $K - \{(l' \to r')\}$ an Position ϵ sein. Sei $\mathcal{U}'' := \mathcal{U} \backslash \mathcal{U}'$. Wegen $t > t'$ sind die Redexe in $\mathcal{U}$ paarweise unkritisch. Außerdem liegen die zu $(l' \to r')$ an ϵ unkritischen Redexe, d.h. die in $\mathcal{U}''$, tiefer in t als die kritischen in $\mathcal{U}'$. Daher bleiben die Redexe in $\mathcal{U}'$ nach Kontraktion der Redexe in $\mathcal{U}''$ erhalten und der Übergang $t > t'$ läßt sich in zwei Schritte aufteilen:

   $$t >^{\mathcal{U}''} t'' >^{\mathcal{U}'} t'$$

Da R nach Voraussetzung linkslinear ist tritt jede freie Variable $F \in fv(l)$ nur einmal in l auf. Deswegen gibt es eine Substitution θ' mit $\theta >^{\mathcal{U}''} \theta'$ und $t'' = \theta'(l)\downarrow_\beta$.

i. $\mathcal{U}'$ sei nicht leer und *total geordnet* durch die Präfixordnung auf den Positionen der Redexe. Dann folgt $t' = \theta'(r')\downarrow_\beta$ aus der Semiorthogonalität. Aus der Induktionshypothese folgt $\theta' > \theta^*$ und somit
$$t' = \theta'(r')\downarrow_\beta > \theta^*(r')\downarrow_\beta = t^*$$
nach Lemma 2.2.7, Lemma 2.2.11 und der Definition von t^*.

ii. Wenn die Positionen von zwei Redexen u_1, u_2 in $\mathcal{U}'$ unvergleichbar (d.h. parallel) sind, dann verändert die Kontraktion dieser beiden Redexe den Term nicht, wegen der Semiorthogonalität und Lemma 6.2.5. Diese beiden Redexe können daher aus $\mathcal{U}'$ entfernt werden, ohne daß sich etwas ändert.
Sei also $\mathcal{V} := \mathcal{U}'\backslash\{u_1, u_2\}$. Dann gilt $t'' >^{\mathcal{V}} t'$. Dieser Vorgang wird so lange wiederholt, bis es keine parallelen Redexe mehr gibt. Dadurch erhält man eine *total geordnete* (möglicherweise leere) Teilmenge $\mathcal{V}' \subseteq \mathcal{U}'$ mit $t'' >^{\mathcal{V}'} t'$. Es gibt wieder zwei Fälle.

A. $\mathcal{V}' = \emptyset$. Daraus folgt $t' = t''$. Aus der Induktionshypothese folgt $\theta' > \theta^*$ und somit folgt mit Lemma 2.2.11
$$t' = \theta'(l')\downarrow_\beta > \theta^*(r')\downarrow_\beta = t^*$$

B. $\mathcal{V}' \neq \emptyset$. Dann gilt
$$t' = \theta'(r')\downarrow_\beta > \theta^*(r')\downarrow_\beta = t^*$$
nach Fall 2(c)i.

Somit ist die Behauptung $t' > t^*$ für jeden Fall gezeigt. □

Korollar 6.2.8 *Sei R ein semiorthogonales HRS und $>$ die von R induzierte parallele bottom-up Reduktion. Dann hat $>$ die Diamanteigenschaft.*
Beweis Sei $t > t_1$ und $t > t_2$, dann folgt aus Lemma 6.2.7
$$t_1 > t^* \ \wedge \ t_2 > t^*$$
□

Satz 6.2.9 *Sei R ein semiorthogonales HRS und $\rightarrow$ die von R induzierte Relation auf den λ-Termen. Dann ist $\rightarrow$ konfluent.*
Beweis direkt aus Korollar 6.2.8, Lemma 2.2.10 und Lemma 2.1.8. □

Beispiel 6.2.10 Das schon mehrfach betrachtete HRS aus Beispiel 1.2.6 mit den Regeln

$$\begin{array}{rcrcl} \beta & : & app(abs(\lambda x.F(x)), S) & \rightarrow & F(S) \\ \eta & : & abs(\lambda x.app(H, x)) & \rightarrow & H \end{array}$$

hat die kritischen Paare

$$\langle app(H, S), app(H, S) \rangle$$

und

$$\langle abs(\lambda x.F(x)), abs(\lambda x.F(x)) \rangle$$

Es ist daher schwach orthogonal, aber nicht orthogonal. Da andere Überlappungen von Regeln als jene, die zu diesen kritischen Paaren führen nicht auftreten gibt es hier keinen kritischen Turm der Höhe zwei oder größer. Daher ist dieses HRS sogar semiorthogonal. Seine Konfluenz folgt somit aus Satz 6.2.9.

Kapitel 7

Kombinatorische Reduktionssysteme

Die Ergebnisse des letzten Kapitels machen deutlich, daß es viel schwieriger ist, Kriterien für die Konfluenz von nicht terminierenden HRS zu finden, als für terminierende HRS. Dies wird insbesondere klar, wenn man die komplizierten Voraussetzungen wie $\twoheadrightarrow\!\!\!\!\!\!+$-Substitutionsstabilität und Semiorthogonalität betrachtet, die zum Beweis der entsprechenden Sätze benötigt werden.

Die bisher behandelten HRS sind jedoch nicht die einzige mögliche Verallgemeinerung der Termersetzungssysteme erster Ordnung. Ein anderer Ansatz sind die von Klop [11] eingeführten kombinatorischen Reduktionssysteme (kurz: CRS). Für diese liegen ebenfalls einige Ergebnisse über Konfluenz vor. Im diesem Kapitel werden die CRS kurz vorgestellt und mit den HRS verglichen.

Kombinatorische Reduktionssysteme (CRS), die zuerst von Klop [11] definiert wurden, sind eine Verallgemeinerung der Termersetzungssysteme erster Ordnung (TRS). Im Gegensatz zu den HRS, die den einfach typisierten λ-Kalkül als Metasprache verwenden, ist die Metasprache bei den CRS eine polyadische Version des $\underline{\lambda}$-Kalküls (siehe [11], bzw. [14]). Im folgenden wird diese Metasprache jedoch nicht explizit definiert, sondern nur implizit angenommen.

7.1 CRS-Terme

Ein CRS ist ein Paar $(\mathcal{A}, \mathcal{R})$, wobei $\mathcal{A}$ das Alphabet der Symbole und $\mathcal{R}$ die Menge der Termersetzungsregeln ist. Bei CRS wird zwischen Termen und Metatermen unterschieden. Die linken und rechten Seiten von Termersetzungsregeln sind Metaterme, doch die von einer Menge $\mathcal{R}$ von Termersetzungsregeln induzierte Relation ist eine Relation auf der Menge der Terme.

Definition 7.1.1 (Symbole) Das Alphabet $\mathcal{A}$ eines CRS $(\mathcal{A}, \mathcal{R})$ besteht aus

- Symbolen für die Variablen $x, y, z, \ldots$
- Dem Abstraktionsoperator, der als [_]_ geschrieben wird
- Symbolen für Operatoren mit einer festen Stelligkeit: $F, G, H, \ldots$
- Symbolen für Metavariablen mit einer festen Stelligkeit: $Z, Z_0, Z_1, \ldots$

Definition 7.1.2 (Metaterme) Die Menge **MTerms** der Metaterme ist die kleinste Menge für die gilt:

1. Für alle Variablen x gilt $x \in MTerms$
2. $[x]t \in MTerms$ für alle Variablen x und $t \in MTerms$
3. $F(t_1, \ldots, t_n) \in MTerms$ für einen n-stelligen Operator F und $t_1, \ldots, t_n \in MTerms$.
4. $Z(t_1, \ldots, t_n) \in MTerms$ für eine n-stellige Metavariable Z und $t_1, \ldots, t_n \in MTerms$.

Die Menge **Terms** der Terme ist die Menge der Metaterme, die keine Metavariablen enthalten. Positionen, Kontexte und Gültigkeitsbereiche gebundener Variablen sind analog zu dem λ-Termen definiert.

Es ist wichtig zu bemerken, daß es im Gegensatz zu λ-Termen bei CRS Termen keine allgemeine Applikation eines Terms auf einen anderen Term gibt. Es gibt nur die Applikation von Variablen, Metavariablen und Funktionskonstanten auf Metaterme.

Beispiel 7.1.3 Die Syntax des λ-Kalküls kann durch CRS-Terme dargestellt werden. Seien ein unärer Abstraktionsoperator λ und ein binärer Applikationsoperator @ gegeben.

λ-Term	CRS-Term
$\lambda x.x$	$\lambda([x]x)$
$(x\ y)$	$@(x, y)$
$\lambda x.(y\ x)$	$\lambda([x]@(y, x))$
$(\lambda x.x)y$	$@(\lambda([x]x), y)$

Im Gegensatz zu diesen Definitionen wird bei HRS kein Unterschied zwischen Variablen und Metavariablen gemacht. Sei F hier eine Funktionskonstante. Eine Variable x ist in einem Term $F(x)$ eine normale Variable, in einer linken bzw. rechten Seite einer Termersetzungsregel $F(x)$ dagegen eine Metavariable.

Bei CRS würde ersteres als $F(x)$ und letzteres als $F(Z)$ geschrieben. Die Metavariablen in CRS-Termersetzungsregeln entsprechen den freien Variablen in den HRS-Termersetzungsregeln.

7.2 CRS-Termersetzungsregeln

Definition 7.2.1 (CRS-Regeln) In einem CRS muß eine Termersetzungsregel $(l \to r)$ folgende Bedingungen erfüllen:

1. l und r sind Metaterme.
2. Das Kopf-Symbol von l ist ein Operatorsymbol.
3. Alle Metavariablen die in r vorkommen kommen auch in l vor.
4. Eine Metavariable Z kommt in l nur in der Form $Z(x_1, \ldots, x_n)$ vor, wobei $x_1, \ldots, x_n$ untereinander verschiedene gebundene Variablen sind.

Diese Bedingungen entsprechen den Bedingungen an die Termersetzungsregeln bei HRS. Die letzte dieser Bedingungen entspricht der Bedingung bei den HRS, daß die linken Seiten der Termersetzungsregeln Patterns sein müssen.

Beispiel 7.2.2 Unter Verwendung der Definitionen aus Beispiel 7.1.3 kann die β-Reduktion des λ-Kalküls $(\lambda x.M)N \to M[x := N]$ auch als CRS-Regel

$$@(\lambda([x]Z(x)), Z') \to Z(Z')$$

geschrieben werden.

7.3 CRS-Termersetzung

Im folgenden wird beschrieben, wie ein CRS eine Relation auf der Menge der CRS-Terme induziert.

Definition 7.3.1 (Ersetzungen) Eine n-stellige Ersetzung ist ein Term der Form

$$\underline{\lambda}(x_1, \ldots, x_n).s$$

wobei s ein Term, $\underline{\lambda}$ ein "Metalambda" und $(x_1, \ldots, x_n)$ n verschiedene Variablen sind, die durch $\underline{\lambda}$ gebunden werden.

So eine Ersetzung kann auf ein n-Tupel von Termen angewandt werden, was zu dem Term

$$(\underline{\lambda}(x_1, \ldots, x_n).s)(t_1, \ldots, t_n) = s[x_1 := t_1, \ldots, x_n := t_n]$$

führt.

Definition 7.3.2 (Zuweisungen) Eine Zuweisung ist eine Abbildung, die jeder n-stelligen Metavariable eine Ersetzung zuordnet, d.h. für eine n-stellige Metavariable Z und eine Zuweisung σ gilt

$$\sigma(Z) = \underline{\lambda}(x_1, \ldots, x_n).s$$

mit einem Term s und Variablen $x_1, \ldots, x_n$.

In Analogie zu den Substitutionen bei den HRS wird diese Abbildung σ zu einer Abbildung von der Menge der Metaterme in die Menge der Terme fortgesetzt.

$$\begin{array}{rcl} x^\sigma &=& x \\ ([x]t)^\sigma &=& [x]t^\sigma \\ (F(t_1, \ldots, t_n))^\sigma &=& F(t_1^\sigma, \ldots, t_n^\sigma) \\ (Z(t_1, \ldots, t_n))^\sigma &=& \sigma(Z)(t_1^\sigma, \ldots, t_n^\sigma) \end{array}$$

In Analogie zu Barendregts Konvention wird dabei angenommen, daß die gebundenen Variablen stets automatisch umbenannt werden, wenn sonst durch Anwendung einer Zuweisung neue ungewollte Bindungen entstehen würden.

Definition 7.3.3 (CRS-Termersetzung) Ähnlich wie bei den HRS induziert ein CRS eine Relation $\rightarrow$ auf der Menge der Terme wie folgt:

$$\begin{array}{rcl} s \rightarrow t &:\Leftrightarrow& \exists(l \rightarrow r), \text{Kontext C mit } s = C[l^\sigma] \ \wedge\ t = C[r^\sigma] \\ &:\Leftrightarrow& \exists(l \rightarrow r), \text{Position p in s mit } s/p = l^\sigma \ \wedge\ t = s[r^\sigma]_p \end{array}$$

Es ist wichtig, zu bemerken, daß bei den CRS im Gegensatz zu den HRS keine β-Reduktion der instanziierten linken und rechten Seiten von Termersetzungsregeln stattfindet. Bei der Anwendung von Zuweisungen auf Metaterme werden zwar Argumente in ihren entsprechenden Platz in Funktionen eingesetzt, jedoch werden, anders als bei der β-Reduktion, dabei entstandene neue Redexe nicht weiter kontrahiert.

Der so entstandene Unterschied wird am folgenden Beispiel deutlich. Im Unterschied zur bisherigen Konvention sollen ab jetzt, wie in CRS-Termen, die Symbole $F, G, H, \ldots$ auch in λ-Termen für Konstanten stehen, und auch freie Variablen mit Kleinbuchstaben bezeichnet werden.

Beispiel 7.3.4 Sei ein HRS mit der Regel $F(\lambda y.z(\lambda x.yx)) \rightarrow z(\lambda x.x)$ und eine Substitution $\sigma : z \mapsto \lambda u.uK$ gegeben. Dann gilt

$$\begin{array}{rcl} F(\lambda y.yK) &=& F(\lambda y.(\lambda x.yx)K)\downarrow_\beta \\ &=& F(\lambda y.(\lambda u.uK)(\lambda x.yx))\downarrow_\beta \\ &=& \sigma(F(\lambda y.z(\lambda x.yx)))\downarrow_\beta \\ &\rightarrow& \sigma(z(\lambda x.x))\downarrow_\beta \\ &=& (\lambda u.uK)\lambda x.x\downarrow_\beta \\ &=& (\lambda x.x)K\downarrow_\beta \\ &=& K \end{array}$$

Das entsprechende CRS mit der Regel $F([y]Z(y)) \to Z([x]x)$ und der Zuweisung $\sigma : Z \mapsto \underline{\lambda}(u).@(u, K)$ verhält sich etwas anders.

$$\begin{array}{rcl} F([y]@(y,K)) & = & (F([y]Z(y)))^\sigma \\ & \to & (Z([x]x))^\sigma \\ & = & @([x]x, K) \end{array}$$

Um obiges HRS vollständig durch ein CRS zu simulieren ist es daher nötig eine weitere CRS-Regel für den β-Reduktionsschritt hinzuzunehmen, und zwar

$$@([x]x, K) \to K$$

7.4 Simulation von HRS durch CRS und umgekehrt

HRS und CRS sind äquivalent in dem Sinne, daß man eine Bijektion zwischen CRS-Termen und λ-Termen angeben, und ein gegebenes HRS durch ein CRS simulieren kann und umgekehrt (siehe dazu [13]). Durch diese Simulationen ist es möglich bestimmte Kriterien für die Konfluenz, die für CRS bewiesen sind auf HRS zu übertragen und umgekehrt.

Die Simulation von CRS durch HRS ist relativ unproblematisch (siehe [13]). Insbesondere ist es möglich, zu einem orthogonalen CRS ein orthogonales HRS zu finden, das dieses simuliert. Da die Konfluenz orthogonaler HRS bewiesen ist [3], sind auch orthogonale CRS konfluent.

Wie das Beispiel 7.3.4 zeigt ist die Simulation von HRS durch CRS nicht so einfach. Es ist dazu nötig eine explizite CRS-Regel für die β-Reduktion hinzuzunehmen.

$$@([x]Z(x), Z') \to_\beta Z(Z')$$

Wie Vincent van Oostrom und Femke van Raamsdonk [13] gezeigt haben kann ein HRS durch ein CRS mit dieser expliziten β-Reduktionsregel simuliert werden. Dabei werden die HRS-Regeln in CRS-Regeln übersetzt und die β-Reduktionsregel hinzugefügt. Ein HRS-Reduktionsschritt wird dann simuliert durch einen CRS-Reduktionsschritt, gefolgt von endlich vielen Anwendungen der expliziten β-Reduktionsregel.

Hier wird der Übergang von einem HRS zum simulierenden CRS, sowie die Überführung von λ-Termen in CRS-Terme kurz skizziert. Für eine ausführlichere Darstellung und den Beweis der Korrektheit und Vollständigkeit dieser Transformation wird auf [13] verwiesen.

Definition 7.4.1 (Transformation von λ-Termen in CRS-Terme)
Die Transformation eines λ-Terms t in langer $\beta\eta$-Normalform in den entsprechenden CRS-Term $< t >$ wird induktiv wie folgt definiert.

- $< \lambda x_1, \ldots, x_m.\ x(t_1, \ldots, t_n) > = [x_1] \ldots [x_m] @_n(x, < t_1 >, \ldots, < t_n >)$
- $< \lambda x_1, \ldots, x_m.\ F(t_1, \ldots, t_n) > = [x_1] \ldots [x_m] F(< t_1 >, \ldots, < t_n >)$

Dabei ist die n-stellige Applikation $@_n$ induktiv definiert durch

$$\begin{aligned} @_1(t, t_1) &= @(t, t_1) \\ @_{n+1}(t, t_1, \ldots, t_n) &= @(@_n(t, t_1, \ldots, t_{n-1}), t_n) \end{aligned}$$

Definition 7.4.2 (Transformation eines HRS in ein CRS)
Die Übersetzung einer HRS-Regel $(l \rightarrow r)$ in die entsprechende CRS-Regel $< l \rightarrow r >$ ist definiert als $< l \rightarrow r > := < l > \rightarrow < r >$, wobei $< l >$ und $< r >$ wie folgt definiert sind.

1. Für die Übersetzung einer linken Seite l gilt
   - $< \lambda x_1, \ldots, x_m . x(t_1, \ldots, t_n) > = [x_1, \ldots, x_m] @_n(x, < t_1 >, \ldots, < t_n >)$, falls $x \in bv(l)$.
   - $< \lambda x_1, \ldots, x_m . z(t_1, \ldots, t_n) > = [x_1, \ldots, x_m] Z(t_1{\downarrow_\eta}, \ldots, t_n{\downarrow_\eta})$, falls $z \in fv(l)$. (Da l ein Pattern ist sind $t_1, \ldots, t_n$ η-äquivalent zu verschiedenen gebundenen Variablen).
   - $< \lambda x_1, \ldots, x_m . F(t_1, \ldots, t_n) > = [x_1, \ldots, x_m] F(< t_1 >, \ldots, < t_n >)$.
2. Für die Übersetzung einer rechten Seite r gilt
   - $< \lambda x_1, \ldots, x_m . x(t_1, \ldots, t_n) > = [x_1, \ldots, x_m] @_n(x, < t_1 >, \ldots, < t_n >)$, falls $x \in bv(r)$.
   - $< \lambda x_1, \ldots, x_m . z(t_1, \ldots, t_n) > = [x_1, \ldots, x_m] Z(< t_1 >, \ldots, < t_n >)$, falls $z \in fv(r)$.
   - $< \lambda x_1, \ldots, x_m . F(t_1, \ldots, t_n) > = [x_1, \ldots, x_m] F(< t_1 >, \ldots, < t_n >)$.

Bei der Transformation eines HRS in ein CRS transformiert man alle seine Regeln wie gerade beschrieben und fügt dann noch die β-Reduktionsregel hinzu.

Da die explizite β-Reduktionsregel mit keiner anderen CRS-Regel überlappt ist es möglich die Eigenschaft der Orthogonalität beim Übergang von einem HRS zum simulierenden CRS zu erhalten. Der Beweis der Konfluenz orthogonaler CRS läßt sich somit auf HRS übertragen (siehe [13]).

Für CRS liegen jedoch noch weitergehende Ergebnisse über Konfluenz vor. Femke van Raamsdonk [12] bewies, daß sogar schwach orthogonale CRS konfluent sind. Diese Aussage ist dagegen für HRS noch unbewiesen.

Leider läßt sich die Methode der Simulation von HRS durch CRS nicht dazu verwenden, dieses Ergebnis auf HRS zu übertragen. Das liegt daran, daß im Gegensatz zur Orthogonalität die schwache Orthogonalität eines HRS bei seiner Simulation durch ein CRS nicht erhalten bleibt.

Beispiel 7.4.3 Sei ein HRS R mit folgenden zwei Regeln gegeben.

1. $F(\lambda y.z(\lambda x.yx)) \rightarrow z(\lambda x.x)$

2. $F(\lambda y.yK) \rightarrow K$

Diese beiden Regeln überlappen. Mit der Substitution $\theta : z \mapsto \lambda u.uK$ ergibt sich das kritische Paar $\langle K, K \rangle$. Da R außerdem linkslinear ist folgt daraus die schwache Orthogonalität.

Mit den Transformationsregeln aus Definition 7.4.2 erhält man das simulierende CRS mit den Regeln

1. $F([y]Z(y)) \rightarrow Z([x]x)$

2. $F([y]@(y, K)) \rightarrow K$

3. $@([x]Z(x), Z') \rightarrow_\beta Z(Z')$

wobei letzteres die explizite β-Reduktionsregel ist. Wieder überlappen die ersten beiden Regeln. Mit der Zuweisung $\sigma : Z \rightarrow \underline{\lambda}.@(x, K)$ ergibt sich das kritische Paar $\langle @([x]x, K), K \rangle$. Somit ist das simulierende CRS nicht schwach orthogonal. Erst die Anwendung der expliziten β-Reduktionsregel auf den ersten Teil des kritischen Paares $@([x]x, K)$ würde den Term K ergeben.

Kapitel 8

Ausblick

Aus den bisher gezeigten Ergebnissen wird deutlich, daß es bedeutend schwieriger ist, Kriterien für die Konfluenz von nicht terminierenden Termersetzungssystemen zu finden, als für die Konfluenz von terminierenden. Dies liegt daran, daß für terminierende Termersetzungssysteme die Konfluenz aus der lokalen Konfluenz folgt, die bedeutend leichter nachzuweisen ist. Unter Verwendung dieser Eigenschaft gibt der Satz 3.0.16 aus Kapitel 3 ein notwendiges und hinreichendes Kriterium für die Konfluenz terminierender HRS. Für allgemeine HRS folgt aber daraus nur ein hinreichendes Kriterium für die Konfluenz.

Zwei weitere hinreichende Kriterien für die Konfluenz von HRS wurden in Kapitel 6 gezeigt. Diese Kriterien sind jedoch ebenfalls nicht notwendig.

8.1 Offene Fragen

Die Kriterien für die Konfluenz von nicht terminierenden HRS sind wegen ihren komplizierten und stark einschränkenden Voraussetzungen noch zu schwach. Ein etwas weitergehendes Ergebnis über die Konfluenz schwach orthogonaler Systeme ist zwar für die kombinatorischen Reduktionssysteme schon bewiesen, ein entsprechendes Ergebnis für die HRS steht aber noch aus. In Abschnitt 6.2 wurde gezeigt, daß die schwache Orthogonalität aus der Semiorthogonalität folgt. Es ist möglich, aber nicht bewiesen, daß auch die Umkehrung gilt. In diesem Fall würde die Konfluenz schwach orthogonaler HRS aus der in Abschnitt 6.2 gezeigten Konfluenz semiorthogonaler HRS folgen.

Es ist wünschenswert, noch bessere Kriterien für die Konfluenz von HRS zu finden. Möglicherweise läßt sich eine Variante des Ergebnisses von Huet über die Konfluenz parallel abgeschlossener Termersetzungssysteme erster Ordnung auch ohne die zusätzliche Voraussetzung der $-\!\!\!\!+\!\!\!\!+\!\!\!\!\rightarrow$-Substitutionsstabilität auf HRS verallgemeinern.

8.2 Erweiterungen

Eine weitere interessante Frage ist, welche Eigenschaften nicht terminierende HRS bei der in Kapitel 4 untersuchten Konfluenz modulo Gleichungssystemen haben.

In den Kapiteln über Konfluenz modulo Gleichungssystemen, Gleichheit modulo Termersetzungssystemen und Konfluenz nicht-noetherscher Termersetzungssysteme wurde stets vorausgesetzt, daß die betrachteten HRS R linkslinear sind. Es bleibt zu untersuchen, welche ähnlichen (wahrscheinlich zumeist schwächeren) Eigenschaften für HRS gelten, die diese Einschränkung nicht erfüllen.

Schließlich ist es eine wichtige Frage, inwieweit sich die Einschränkung der linken Seiten der Termersetzungsregeln auf Patterns abschwächen läßt. Insbesondere wäre es wichtig, herauszufinden welche der hier gezeigten Ergebnisse (z.B. über die lokale Konfluenz (Kap.3) oder die Äquivalenz von Termersetzung und Gleichheit modulo eines HRS (Kap.5)) auf welche Termersetzungssysteme mit allgemeineren linken Seiten übertragbar sind.

Abbildungsverzeichnis

Literaturverzeichnis

[1] G. Huet. Confluent reductions: Abstract properties and applications to term rewriting systems. J. ACM, 27:797–821, 1980.

[2] Tobias Nipkow. Higher-Order critical pairs. In *Proceedings, Sixth Annual IEEE Symposium on Logic in Computer Science*, pages 342–349, Amsterdam, The Netherlands, July 15–18 1991. IEEE Computer Society Press.

[3] Tobias Nipkow. *Orthogonal higher-order rewrite systems are confluent.* In M.A. Bezem and Jan Friso Groote, editors, *Proc. Int. Conf. Typed Lambda Calculi and Applications*, pages 306–317. LNCS 664, 1993.

[4] D. Miller. A logic programming language with lambda-abstraction, function variables and simple unification. In P. Schroeder-Heister, editor, *Extensions of Logic Programming*, pages 253–281. LNCS 475, 1991.

[5] J. Hindley and J. P. Seldin. *Introduction to Combinators and λ-Calculus.* Cambridge University Press, 1986.

[6] Masako Takahashi. *λ-Calculi with Conditional Rules.* In M.A. Bezem and Jan Friso Groote, editors, *Proc. Int. Conf. Typed Lambda Calculi and Applications*, pages 406–417. LNCS 664, 1993.

[7] N. Dershowitz and J. P. Jouannaud. *Rewrite systems.* In J. van Leeuwen, editor, *Formal Models and Semantics, Handbook of Theoretical Computer Science, Vol. B*, pages 243–320. Elsevier - The MIT Press, 1990.

[8] W. D. Goldfarb. *The undecidability of the second-order unification problem.* TCS, 13:225–230, 1981.

[9] T. Nipkow. *Functional unification of higher-order patterns.* Technical report, TU München, Institut für Informatik, 1992.

[10] D. Knuth and P. Bendix. *Simple word problems in universal algebra.* In J. Leech, editor, Computational Problems in Abstract Algebra, pages 263–297. Pergamon Press, 1970.

[11] J. W. Klop. *Combinatory Reduction Systems.* Mathmatical Centre Tracts 127. Mathematisch Centrum, Amsterdam, 1980.

[12] F. van Raamsdonk: *A simple proof of confluence for weakly orthogonal combinatory reduction systems.* Technical Report CS-R9234, CWI, Amsterdam 1992.

[13] Vincent van Oostrom and Femke van Raamsdonk. *Comparing Combinatory Reduction Systems and Higher-Order Rewrite Systems.* 1991.

[14] H.P. Barendregt. *The Lambda Calculus, its Syntax and Semantics*, volume 103 of *Studies in Logic and the Foundations of Mathmatics.* North-Holland Publishing Company, revised edition, 1984. (Second Printing 1985).

[15] L.C. Paulson. Isabelle: The next 700 theorem provers. In P. Odifreddi, editor, *Logic and Computer Science*, pages 361–385. Academic Press, 1990.

[16] G. Nadathur and D. Miller. *An overview of λProlog.* In R. A. Kowalski and K.A. Bowen, editors, *Proc. 5th Int. Logic Programming Conference*, pages 810–827. MIT Press, 1988.

[17] A. Felty. *A logic-programming approach to implementing higher-order term rewriting.* In L. H. Eriksson, L. Hallnäs, and P. Schroeder-Heister, editors, *Extensions of Logic Programming, Proc. 2nd Int. Workshop*, pages 135–158. LNCS 596, 1992.

Anhang A

Verzeichnis der wichtigsten Sätze

Satz 3.0.16 31
Sei ein HRS R gegeben. Die von R induzierte Relation $\to$ ist genau dann lokal konfluent, wenn für alle kritischen Paare $\langle u_1, u_2 \rangle$ von R gilt: $u_1 \downarrow u_2$.

Satz 3.0.17 32
Die Eigenschaft der Konfluenz ist für terminierende HRS entscheidbar.

Satz 3.0.19 32
Sei ein HRS R gegeben. Wenn die durch R induzierte Relation $\to$ terminierend ist und für alle kritischen Paare $\langle u_1, u_2 \rangle$ von R $u_1 \downarrow u_2$ gilt, dann ist $\to$ konfluent und jeder Term hat eine eindeutige $\to$-Normalform.

Satz 4.3.1 41
*Sei $\langle R, \mathcal{E} \rangle$ eine Gleichungstheorie, R linkslinear und $(\to \cdot \sim)$ terminierend. Es stehe $\hat{t}$ für eine beliebige $\to$-Normalform eines Terms t. Dann gilt:
Die Gleichungstheorie $\langle R, \mathcal{E} \rangle$ ist genau dann konfluent, wenn für alle kritischen Paare $\langle u_1, u_2 \rangle$ von $\langle R, \mathcal{E} \rangle$ gilt: $\widehat{u_1} \sim \widehat{u_2}$.*

Satz 5.1.10 49
Sei R ein HRS und s, t λ-Terme. Dann gilt

$$s =_R t \quad \Leftrightarrow \quad s\updownarrow_\beta^\eta \overset{*}{\underset{R}{\leftrightarrow}} t\updownarrow_\beta^\eta$$

Satz 5.2.2 50
Sei R ein HRS mit der zulässigen Erweiterung R'. Es gelte $R' = \mathcal{E} \dot{\cup} D$ (disjunkte Vereinigung) so daß folgende Voraussetzungen erfüllt sind:

- *$\mathcal{E}$ ist ein symmetrisches HRS, $\vdash\!\dashv_{\mathcal{E}}$ die durch $\mathcal{E}$ induzierte Relation, $\sim := \overset{*}{\vdash\!\dashv}_{\mathcal{E}}$ und $=_{\mathcal{E}}$ die durch $\mathcal{E}$ induzierte Gleichheit.*
- *D ist ein linkslineares HRS, $\rightarrow_D$ die durch D induzierte Relation und $\hat{t}$ eine beliebige $\rightarrow_D$-Normalform eines λ-Terms t.*
- *$(\rightarrow_D \cdot \sim)$ terminiert.*
- *Für alle kritischen Paare $\langle u_1, u_2 \rangle$ der so entstandenen Gleichungstheorie $\langle D, \mathcal{E} \rangle$ gilt $\widehat{u_1} =_{\mathcal{E}} \widehat{u_2}$.*

Dann gilt folgende Äquivalenz:

$$s =_R t \quad \Leftrightarrow \quad \widehat{s\updownarrow^{\eta}_{\beta}} =_{\mathcal{E}} \widehat{t\updownarrow^{\eta}_{\beta}}$$

Satz 6.1.8 62
Sei R ein HRS, das $\mathrel{-\!\!\|\!\!\rightarrow}$-substitutionsstabil, parallel abgeschlossen und linkslinear ist. Dann ist die von R induzierte Relation $\rightarrow_R$ konfluent.

Satz 6.2.9 69
Sei R ein semiorthogonales HRS und $\rightarrow$ die von R induzierte Relation auf den λ-Termen. Dann ist $\rightarrow$ konfluent.