Analyzing GÖDEL's T via expanded head reduction trees

Arnold Beckmann*
Mathematical Institute, University of Oxford
24-29 St. Giles', Oxford OX1 3LB, UK†

Andreas Weiermann‡
Institut für mathematische Logik und Grundlagenforschung
der Westfälischen Wilhelms-Universität Münster
Einsteinstr. 62, D-48149 Münster, Germany§

February 26, 1999

Abstract

Inspired from BUCHHOLZ' ordinal analysis of ID_1 and BECKMANN's analysis of the simple typed λ-calculus we classify the derivation lengths for GÖDEL's system T in the λ-formulation (where the η-rule is included).

1 Introduction

In this paper we develop a perspicuous method for classifying the derivation lengths of GÖDEL's T. Following ideas from [Be98] we assign canonically to each term $t \in T$ an expanded head reduction tree. The size of this tree, if it is finite, yields a nontrivial bound on the maximal length of a reduction chain starting with t, since the expanded head reduction trees represent worst case reductions. Using ideas from infinitary proof theory we show that it is indeed possible to define a finite expanded head reduction tree for any term of T. For this purpose we enlarge the concept of expanded head reduction trees by a cut rule and an appropriate miniaturization of BUCHHOLZ' Ω-rule (for dealing with terms containing recursors). The embedding and cut elimination procedure is carried out by adapting BUCHHOLZ' treatment of ID_1 (cf. [Bu80]). To obtain optimal complexity bounds even for the fragments of T we utilize a system $\mathcal{T}$ of formal ordinal terms for the ordinals less than ε_0 and an appropriate collapsing function $\mathcal{D} : \mathcal{T} \to \omega$. To obtain an unnested recursive definition of $\mathcal{D}$ we utilize crucial properties of the theory of the ψ function which is developed, for example, in [W98].

*Partially supported by the Deutschen Akademie der Naturforscher Leopoldina grant #BMBF-LPD 9801-7 with funds from the Bundesministeriums für Bildung, Wissenschaft, Forschung und Technologie.

†email: Arnold.Beckmann@math.uni-muenster.de

‡Supported by a Heisenberg grant of the Deutsche Forschungsgemeinschaft.

§email: weierma@math.uni-muenster.de

Compared with prior treatments of classifying the T-derivation lengths (cf., e.g., [W98, WW98]) the method described in this paper has the advantage that the ordinals assigned to the terms of T are assigned in a more genuine and intrinsic way.

2 Expanded head reduction trees

The derivation length $\mathrm{d}(r)$ of a term r is the longest possible reduction sequence starting from r:

$$\mathrm{d}(r) := \max\{k : \exists s \in \mathrm{T}(\mathcal{V}), r \longrightarrow^k s\}.$$

In case of simple typed λ-calculus it is shown in [Be98] that computing the expanded head reduction tree of r leads to estimations on $\mathrm{d}(r)$. Here we will extend this approach to GÖDEL's T. To this end we first have to fix what the head redex of a term is. Of course the presence of the recursor R makes thing much more complicated than in the case of simple typed λ-calculus. The head redex can occur deep inside the term. E.g. let $b := \lambda x.\lambda y.\,\mathrm{S}\,y$, then the head redex of $v := \mathrm{R}(\mathrm{R}((\lambda x.\,\mathrm{S}\,x)t)bc)de$ is $(\lambda x.\,\mathrm{S}\,x)t$, so v reduces with head reductions in the following way:

$$\begin{aligned} v &\longrightarrow^1 \mathrm{R}(\mathrm{R}(\mathrm{S}\,t)bc)de \longrightarrow^1 \mathrm{R}(bt(\mathrm{R}\,tbc))de \\ &\longrightarrow^2 \mathrm{R}(\mathrm{S}(\mathrm{R}\,tbc))de \longrightarrow^1 d(\mathrm{R}\,tbc)(\mathrm{R}(\mathrm{R}\,tbc)de) \end{aligned}$$

The terms $\mathrm{T}(\mathcal{V})$ of GÖDEL's T are build up from a set of variables $\mathcal{V}$ (countably many for each type) and the symbols for the recursor R for any type, for zero 0 of type 0 and for the successor S of type $0 \to 0$. We will decompose every term $t \in \mathrm{T}(\mathcal{V})$ into its head redex $\mathrm{redex}(t)$ and the rest $\mathrm{coat}_t(\star)$ which we call coat such that $t = \mathrm{coat}_t(\mathrm{redex}(t))$. Not every head redex is reducible, e.g. if $\mathrm{redex}(t)$ starts with a variable. In this case reductions are only possible in all other terms which occur up to the depth of $\mathrm{redex}(t)$ and these reductions can be considered in parallel. Therefore we collect those terms into a multiset $\mathrm{mat}(t)$ called the material of t. Furthermore we split the redex of t into its characteristic part $\mathrm{rr}(t)$ which is needed to define the expanded head reduction tree.

With $\{\{\dots\}\}$ we indicate multisets, with $\cup$ their union and with $\#$ their cardinality. Let $\mathcal{V}_{\mathrm{R}} := \mathcal{V} \cup \{\mathrm{R}\}$, $\mathcal{V}_{0,\mathrm{S}} := \mathcal{V} \cup \{0, \mathrm{S}\}$ and $\mathcal{V}_{0,\mathrm{S},\mathrm{R}} := \mathcal{V}_{0,\mathrm{S}} \cup \{\mathrm{R}\}$.

Definition 2.1 *We define terms* $\mathrm{redex}(s), \mathrm{rr}(s) \in \mathrm{T}(\mathcal{V})$ *and* $\mathrm{coat}_s(\star) \in \mathrm{T}(\mathcal{V} \cup \{\star\})$ *and a multiset of* $\mathrm{T}(\mathcal{V})$*-terms* $\mathrm{mat}(s)$ *by recursion on* $s \in \mathrm{T}(\mathcal{V})$*.*

s		$\mathrm{rr}(s)$	$\mathrm{redex}(s)$	$\mathrm{coat}_s(\star)$	$\mathrm{mat}(s)$
$x\vec{t}$	$x \in \mathcal{V}_{0,\mathrm{S}}$	x	$x\vec{t}$	$\star$	$\{\{\vec{t}\}\}$
λxr		$(\lambda xr)x$	λxr	$\star$	$\emptyset$
$(\lambda xr)u\vec{t}$		$(\lambda xr)u$	$(\lambda xr)u$	$\star\vec{t}$	$\emptyset$
$\mathrm{R}\,u_1 \dots u_l$	$l \leq 2$	R	$\mathrm{R}\,\vec{u}$	$\star$	$\{\{\vec{u}\}\}$
$\mathrm{R}\,tab\vec{s}$					
$t = 0, \mathrm{S}\,t'$		$\mathrm{R}\,tab$	$\mathrm{R}\,tab$	$\star\vec{s}$	$\emptyset$
$t = x\vec{u},\ x \in \mathcal{V}$		x	$\mathrm{R}\,tab$	$\star\vec{s}$	$\{\{\vec{u}, a, b, \vec{s}\}\}$
$t \neq x\vec{u},\ x \in \mathcal{V}_{0,\mathrm{S}}$		$\mathrm{rr}(t)$	$\mathrm{redex}(t)$	$\mathrm{R}\,\mathrm{coat}_t(\star)ab\vec{s}$	$\mathrm{mat}(t) \cup \{\{a, b, \vec{s}\}\}$

Obviously we have $\mathrm{coat}_t(\mathrm{redex}(t)) = t$, $\mathrm{rr}(t) = \mathrm{rr}(\mathrm{redex}(t))$ and

$$\begin{aligned}
\mathrm{rr}(t) &\in \mathcal{V} \cup \{0, \mathrm{S}, \mathrm{R}\} \cup \{(\lambda xr)s, \mathrm{R}\,0ab, \mathrm{R}(\mathrm{S}\,s)ab \mid a, b, r, s \in \mathrm{T}(\mathcal{V})\} \\
\mathrm{redex}(t) &\in \{\lambda xr, (\lambda xr)s, y\vec{t}, \mathrm{R}(y\vec{t})rs \mid r, s, \vec{t} \in \mathrm{T}(\mathcal{V}), y \in \mathcal{V}_{0,\mathrm{S}}\} \\
&\quad \cup \{\mathrm{R}\,u_1 \dots u_l \mid l \leq 2 \;\&\; \vec{u} \in \mathrm{T}(\mathcal{V})\}
\end{aligned}$$

Definition 2.2 *We inductively define* $\overset{\alpha}{\models} t$ *for* $t \in \mathrm{T}(\mathcal{V})$ *and* $\alpha < \omega$ *if one of the following cases holds:*

($\mathcal{V}_{0,\mathrm{S},\mathrm{R}}$-Rule) $\mathrm{rr}(t) \in \mathcal{V}_{0,\mathrm{S},\mathrm{R}}$ *and there is some* β *such that* $\beta + \#\,\mathrm{mat}(t) \leq \alpha$ *and* $\forall s \in \mathrm{mat}(t) \overset{\beta}{\models} s$*.*

(β-Rule) $\mathrm{rr}(t) = (\lambda xr)s$ *and* $\overset{\beta}{\models} \mathrm{coat}_t(r[x := s])$ *and* $\overset{\beta}{\models} s$ *for some* $\beta < \alpha$*.*

(R 0-Rule) $\mathrm{rr}(t) = \mathrm{R}\,0ab$ *and* $\overset{\beta}{\models} \mathrm{coat}_t(b)$ *and* $\overset{\beta}{\models} a$ *for some* $\beta < \alpha$*.*

(R S-Rule) $\mathrm{rr}(t) = \mathrm{R}(\mathrm{S}\,t')ab$ *and* $\overset{\beta}{\models} \mathrm{coat}_t(at'(\mathrm{R}\,t'ab))$ *for some* $\beta < \alpha$*.*

The β-Rule is well-defined because $\mathrm{redex}(t) = \lambda xr \Rightarrow t = \lambda xr$*. Observe that we have* $\mathrm{redex}(t) = \mathrm{rr}(t)$ *for the* R 0*-Rule and the* R S*-Rule .*

Obviously $\overset{0}{\models} x$ for any variable x and 0, S.

We observe that $\overset{\alpha}{\models} r$ can be viewed as a tree which is generated in a unique way. We call this tree (with the α's stripped off) the expanded head reduction tree of r. We are going to define a number $\#t$ for any term t which computes the number of nodes with conversion in the expanded head reduction tree of that term.

Definition 2.3 *Define* $\#t$ *for* $t \in \mathrm{T}(\mathcal{V})$ *by recursion on* $\overset{\alpha}{\models} t$*. This is well-defined because the expanded head reduction tree is unique.*

1. $\mathrm{rr}(t) \in \mathcal{V} \cup \{0, \mathrm{S}, \mathrm{R}\}$ *then* $\#t := \sum_{s \in \mathrm{mat}(t)} \#s$*.*
2. $\mathrm{rr}(t) = (\lambda xr)s$ *then* $\#t := \#\,\mathrm{coat}_t(r[x := s]) + \#s + 1$*.*
3. $\mathrm{rr}(t) = \mathrm{R}\,0ab$ *then* $\#t := \#\,\mathrm{coat}_t(b) + \#a + 1$*.*
4. $\mathrm{rr}(t) = \mathrm{R}(\mathrm{S}\,t')ab$ *then* $\#t := \#\,\mathrm{coat}_t(at'(\mathrm{R}\,t'ab)) + 1$*.*

Lemma 2.4 *If* $\mathrm{rr}(r) \neq z$ *and* $z \in \mathcal{V}$ *then*

1. $\mathrm{redex}(r[z := s]) = \mathrm{redex}(r)[z := s]$
2. $\mathrm{coat}_{r[z:=s]}(\star) = \mathrm{coat}_r(\star)[z := s]$
3. $\mathrm{mat}(r[z := s]) = \mathrm{mat}(r)[z := s]$

Lemma 2.5 *Assume* $\mathrm{rr}(r) = z \in \mathcal{V} \cup \{\mathrm{R}\}$*. If* $\mathrm{redex}(r) = z\vec{t}$ *then* $\mathrm{redex}(r) = r$*. Otherwise* $\mathrm{redex}(r) = \mathrm{R}(z\vec{t})ab$ *and* $z \in \mathcal{V}$*, thus*

1. $\mathrm{mat}(r) = \mathrm{mat}(\mathrm{coat}_r(z')) \cup \{\{\vec{t}, a, b\}\}$ *for some suitable* z'*.*
2. *if* $s \in \mathrm{T}(\mathcal{V})$ *with* $z \notin \mathrm{fvar}(s)$ *then*

(a) $\mathrm{redex}(r[z:=s]) = \mathrm{redex}(\mathrm{R}(z\vec{t})[z:=s]ab)[z:=s]$

(b) $\mathrm{coat}_{r[z:=s]}(\star) = \mathrm{coat}_r(\mathrm{coat}_{\mathrm{R}(z\vec{t})[z:=s]ab}(\star))[z:=s]$

In order to handle η-reductions we need $\#rx \geq \#r$, then we can compute $\#\lambda x.px = \#px + 1 > \#p$. But in order to obtain $\#rx \geq \#r$ we need a Lemma which comes with a rather technical proof.

Lemma 2.6 *$\#r[z:=u] \geq \#r + \#u$ if $z \in \mathrm{fvar}(r)$.*

Using this we immediately obtain

Lemma 2.7 *1. $\#r[x:=y] = \#r$.*
2. $\#rx \geq \#r$.

Proof. 1. is clear.
For 2. we compute $\#rx \geq \#yx + \#r \geq \#r$ using Lemma 2.6 for the first $\geq$. QED.

Proof of Lemma 2.6. More generally we will prove

$$\forall r,u \in \mathrm{T}(\mathcal{V}) \forall z \in \mathcal{V}\Big(z \text{ occurs exactly once free in } r \text{ and } \#r[z:=u] = k \\ \Rightarrow \#r + \#u \leq \#r[z:=u]\Big)$$

by induction on k. Let k, r, u, z fulfill the premise of this assertion. Define s^* to be $s[z:=u]$ for terms s.

$\mathrm{rr}(r) = (\lambda xs)t$. By Lemma 2.4 we have $\mathrm{redex}(r^*) = \mathrm{redex}(r)^*$ and $\mathrm{coat}_{r^*} = \mathrm{coat}_r^*$, thus $\mathrm{rr}(r^*) = \mathrm{rr}(\mathrm{redex}(r^*)) = \mathrm{rr}((\lambda xs^*)t^*) = (\lambda xs^*)t^*$. Hence

$$\begin{aligned} \#r^* &= \#coat_r(s[x:=t])^* + \#t^* + 1 \\ &\overset{*1}{\geq} \#coat_r(s[x:=t]) + \#t + 1 + \#u = \#r + \#u \end{aligned}$$

where for estimation $*1$ we used the induction hypothesis eventually several times.

Similar are the cases for $\mathrm{rr}(r) = \mathrm{R}\,0ab$, $\mathrm{rr}(r) = \mathrm{R}(\mathrm{S}\,t)ab$ and $\mathrm{rr}(r) = y \in \mathcal{V} \cup \{0, \mathrm{S}, \mathrm{R}\}$ with $y \neq z$.

The case $\mathrm{rr}(r) = z$ needs very much effort. Observe that $\mathrm{rr}(r)$ is the only occurrence of z in r.

- $\mathrm{redex}(u) = y\vec{v}$ with $y \in \mathcal{V}_{0,\mathrm{S}}$, hence $u = y\vec{v}$ by Lemma 2.5. In the case $\mathrm{redex}(r) = z\vec{t}$ Lemma 2.5 shows $r = z\vec{t}$, hence

$$\#r^* = \#y\vec{v}\vec{t} = \sum \#\vec{v} + \sum \#\vec{t} = \#u + \#r.$$

Otherwise $\mathrm{redex}(r) = \mathrm{R}(z\vec{t})ab$ and Lemma 2.5 2. shows

$$\mathrm{redex}(r^*) = \mathrm{R}(y\vec{v}\vec{t})ab$$
$$\mathrm{coat}_{r^*}(\star) = \mathrm{coat}_r(\star)$$

1. $y \in \mathcal{V}$, then Lemma 2.5 1. shows

$$\begin{aligned}\mathrm{mat}(r^*) &= \mathrm{mat}(\mathrm{coat}_r(z')) \cup \{\{\vec{v}, \vec{t}, a, b\}\} \\ &= \mathrm{mat}(r) \cup \mathrm{mat}(u)\end{aligned}$$

   Thus

$$\#r^* = \sum_{v \in \mathrm{mat}(r)} \#v + \sum_{v \in \mathrm{mat}(u)} \#v = \#r + \#u.$$

2. $y = 0$, hence $y\vec{v}\vec{t} = 0$ and we compute

$$\begin{aligned}\#r^* &= \#\,\mathrm{coat}_r(b) + \#a + 1 \overset{i.h.}{\geq} \#\,\mathrm{coat}_r(z') + \#b + \#a + 1 \\ &= \#r + 1 > \#r + \#u\end{aligned}$$

   where the last equation uses Lemma 2.5 1.

3. $y = \mathrm{S}$, hence $y\vec{v}\vec{t} = \mathrm{S}\,v$ and we compute

$$\begin{aligned}\#r^* &= \#\,\mathrm{coat}_r(av(\mathrm{R}\,vab)) + 1 \overset{i.h.}{\geq} \#\,\mathrm{coat}_r(z') + \#a + \#b + \#v + 1 \\ &= \#r + \#u + 1 > \#r + \#u\end{aligned}$$

   and we used the induction hypothesis several times.

- $\mathrm{redex}(u) = \mathrm{R}(y\vec{v})cd$ with $y \in \mathcal{V}_{0,\mathrm{S}}$. If $\mathrm{redex}(r) = z\vec{t}$ then

$$\begin{aligned}\mathrm{redex}(r^*) &= \mathrm{R}(y\vec{v})cd \\ \mathrm{coat}_{r^*}(\star) &= \mathrm{coat}_u(\star)\vec{t},\end{aligned}$$

otherwise $\mathrm{redex}(r) = \mathrm{R}(z\vec{t})ab$, hence

$$\begin{aligned}\mathrm{redex}(r^*) &= \mathrm{R}(y\vec{v})cd \\ \mathrm{coat}_{r^*}(\star) &= \mathrm{coat}_r(\mathrm{R}(\mathrm{coat}_u(\star)\vec{t})ab)\end{aligned}$$

Similar to the previous case we compute

$$\#r^* \geq \#r + \#u.$$

For the following cases we state some abbreviations. If $\mathrm{redex}(r) = z\vec{t}$ then $r = z\vec{t}$ by Lemma 2.5. Let $c_r(\star) := \star$. Otherwise $\mathrm{redex}(r) = \mathrm{R}(z\vec{t})ab$. Let $c_r(\star) := \mathrm{coat}_r(\mathrm{R}\,\star ab)$. In both cases we have using Lemma 2.5

$$\begin{aligned}r &= c_r(z\vec{t}) \\ \mathrm{mat}(r) &= \mathrm{mat}(c_r(z')) \cup \{\{\vec{t}\}\}\end{aligned}$$

- $\mathrm{redex}(u) = \mathrm{R}\,u_1 \ldots u_l$ with $l \leq 2$. Then $u = \mathrm{R}\,u_1 \ldots u_l$ by Lemma 2.5.
  Let $u_1 \ldots u_l\vec{t} =: v_1 \ldots v_m$ for some $m, v_1, \ldots, v_m$. Then

$$\begin{aligned}r^* &= c_r(\mathrm{R}\,\vec{v}) \\ \mathrm{redex}(r^*) &= \mathrm{redex}(\mathrm{R}\,\vec{v}) \\ \mathrm{coat}_{r^*}(\star) &= c_r(\mathrm{coat}_{\mathrm{R}\,\vec{v}}(\star))\end{aligned}$$

We compute with Lemma 2.5 1.

$$\#r + \#u = \#c_r(z') + \sum \#\vec{t} + \sum_{j=1}^{l} \#u_j = \#c_r(z') + \sum_{j=1}^{m} \#v_j \tag{1}$$

We distinguish the cases for m and $\mathrm{redex}(v_1)$.

1. $m \leq 2$, then by Lemma 2.5 $r = z\vec{t}$, $r^* = \mathrm{R}\, v_1 \dots v_m$ and $c_r(\star) = \star$, hence

$$\#r^* = \sum_{j=1}^{m} \#v_j = \#c_r(z') + \sum_{j=1}^{m} \#v_j \overset{(1)}{=} \#r + \#u$$

   For the following cases assume $m \geq 3$.

2. $v_1 = 0$, then $\mathrm{redex}(r^*) = \mathrm{R}\, 0 v_2 v_3$ and $\mathrm{coat}_{r^*}(\star) = c_r(\star v_4 \dots v_m)$. Thus

$$\begin{aligned} \#r^* &= \#c_r(v_3 \dots v_m) + \#v_2 + 1 \\ &\overset{i.h.}{>} \#c_r(z') + \#(v_3 \dots v_m) + \#v_2 \\ &\overset{*2}{\geq} \#c_r(z') + \#(x_3 \dots x_m) + \sum_{j=2}^{m} \#v_j \\ &\geq \#c_r(z') + \sum_{j=1}^{m} \#v_j \overset{(1)}{=} \#r + \#u \end{aligned}$$

   where for estimation $*2$ we used several times the induction hypothesis and $x_3 \dots x_m$ are suitable new variables.

3. $v_1 = \mathrm{S}\, v$, then $\mathrm{redex}(r^*) = \mathrm{R}(\mathrm{S}\, v) v_2 v_3$ and $\mathrm{coat}_{r^*}(\star) = c_r(\star v_4 \dots v_m)$. Hence

$$\begin{aligned} \#r^* &= \#c_r(v_2 v (\mathrm{R}\, v v_2 v_3) v_4 \dots v_m) + 1 \\ &\overset{i.h.}{>} \#c_r(z') + \#v_2 v (\mathrm{R}\, x_1 x_2 x_3) x_4 \dots x_m + \#v + \sum_{j=2}^{m} \#v_j \\ &\overset{*3}{\geq} \#c_r(z') + \sum_{j=1}^{m} \#v_j \overset{(1)}{=} \#r + \#u \end{aligned}$$

   where for estimation $*3$ we observe $\#v = \#\, \mathrm{S}\, v$.

4. $\mathrm{redex}(v_1) = x\vec{w}$ with $x \in \mathcal{V}$, then $v_1 = x\vec{w}$, thus $\mathrm{redex}(r^*) = \mathrm{R}(x\vec{w}) v_2 v_3$ and $\mathrm{coat}_{r^*}(\star) = c_r(\star v_4 \dots v_m)$. Hence

$$\begin{aligned} \#r^* &= \#c_r(z') + \sum_{w \in \mathrm{mat}(v_1)} \#w + \sum_{j=2}^{m} \#v_j \\ &= \#c_r(z') + \sum_{j=1}^{m} \#v_j \overset{(1)}{=} \#r + \#u \end{aligned}$$

5. $\mathrm{redex}(v_1) = \mathrm{R}\, w_1 \dots w_n$ with $n \leq 2$ and $\mathrm{redex}(v_1) = \lambda x s$ not possible because $\mathrm{lev}(v_1) = 0$.

6. $\mathrm{redex}(v_1) = (\lambda xs)t$, then $\mathrm{redex}(r^*) = (\lambda xs)t$ and

$$\mathrm{coat}_{r^*}(\star) = c_r(\mathrm{R}\,\mathrm{coat}_{v_1}(\star)v_2 \ldots v_m).$$

Hence

$$\begin{aligned}
\#r^* &= \#c_r(\mathrm{R}\,\mathrm{coat}_{v_1}(s[x := t])v_2 \ldots v_m) + \#s + 1 \\
&\overset{i.h.}{\geq} \#c_r(z') + \#\,\mathrm{coat}_{v_1}(s[x := t]) + \#s + 1 + \sum_{j=2}^{m} \#v_j \\
&= \#c_r(z') + \sum_{j=1}^{m} \#v_j \overset{(1)}{=} \#r + \#u
\end{aligned}$$

7. $\mathrm{redex}(v_1) = \mathrm{R}\,0cd$, then $\mathrm{redex}(r^*) = \mathrm{R}\,0cd$

$$\mathrm{coat}_{r^*}(\star) = c_r(\mathrm{R}\,\mathrm{coat}_{v_1}(\star)v_2 \ldots v_m).$$

Hence

$$\begin{aligned}
\#r^* &= \#c_r(\mathrm{R}\,\mathrm{coat}_{v_1}(d)v_2 \ldots v_m) + \#c + 1 \\
&\overset{i.h.}{\geq} \#c_r(z') + \#\,\mathrm{coat}_{v_1}(d) + \#c + 1 + \sum_{j=2}^{m} \#v_j \\
&= \#c_r(z') + \sum_{j=1}^{m} \#v_j \overset{(1)}{=} \#r + \#u
\end{aligned}$$

8. $\mathrm{redex}(v_1) = \mathrm{R}(\mathrm{S}\,w)cd$, then $\mathrm{redex}(r^*) = \mathrm{R}(\mathrm{S}\,w)cd$

$$\mathrm{coat}_{r^*}(\star) = c_r(\mathrm{R}\,\mathrm{coat}_{v_1}(\star)v_2 \ldots v_m).$$

Hence

$$\begin{aligned}
\#r^* &= \#c_r(\mathrm{R}\,\mathrm{coat}_{v_1}(cw(\mathrm{R}\,wcd))v_2 \ldots v_m) + 1 \\
&\overset{i.h.}{\geq} \#c_r(z') + \#\,\mathrm{coat}_{v_1}(cw(\mathrm{R}\,wcd)) + 1 + \sum_{j=2}^{m} \#v_j \\
&= \#c_r(z') + \sum_{j=1}^{m} \#v_j \overset{(1)}{=} \#r + \#u
\end{aligned}$$

• $\mathrm{redex}(u) = \lambda xs$ then $u = \lambda xs$ by induction on the definition of $\mathrm{redex}(u)$. If $z\vec{t} = z$ then $r = z$ because $\mathrm{lev}(z) > 0$. Hence

$$\#r^* = \#u = \#r + \#u.$$

Otherwise $z\vec{t} = zv_0\vec{v}$, thus $\mathrm{redex}(r^*) = (\lambda xs)v_0$ and $\mathrm{coat}_{r^*}(\star) = c_r(\star\vec{v})$. Hence

$$\begin{aligned}
\#r^* &= \#c_r(s[x := v_0]\vec{v}) + \#v_0 + 1 \\
&\overset{i.h.}{\geq} \#c_r(z') + \#s[x := v_0] + \#\vec{v} + \#v_0 + 1 \\
&\overset{(i.h.)}{\geq} \#c_r(z') + \#v_0 + \#\vec{v} + \#s + 1 \\
&\overset{*4}{=} \#c_r(z\vec{t}) + \#\lambda xs = \#r + \#u
\end{aligned}$$

With (i.h.) we mean that we eventually used the induction hypothesis and at *4 we used Lemma 2.5.

• $\mathrm{redex}(u) = (\lambda xs)v$, then $\mathrm{redex}(r^*) = (\lambda xs)v$ and $\mathrm{coat}_{r^*}(\star) = c_r(\mathrm{coat}_u(\star)\vec{t})$. Hence

$$\begin{aligned}\#r^* &= \#c_r(\mathrm{coat}_u(s[x := v])\vec{t}) + \#v + 1\\ &\overset{i.h.}{\geq} \#c_r(z\vec{t}) + \#\,\mathrm{coat}_u(s[x := v]) + \#v + 1 = \#r + \#u\end{aligned}$$

• The cases for $\mathrm{redex}(u) = \mathrm{R}\,0cd$ and for $\mathrm{redex}(u) = \mathrm{R}(\mathrm{S}\,v)cd$ are similar to the previous one. QED.

Main Lemma 2.8 $r \longrightarrow^1 s \Rightarrow \#r > \#s$

Proof. More generally we show for r such that z occurs exactly once:

1. $\#r[z := (\lambda xp)q] > \#r[z := p[x := q]]$
2. $\#r[z := \lambda x.px] > \#r[z := p]$ if $x \notin \mathrm{fvar}(p)$
3. $\#r[z := \mathrm{R}\,0ab] > \#r[z := b]$
4. $\#r[z := \mathrm{R}(\mathrm{S}\,t)ab] > \#r[z := at(\mathrm{R}\,tab)]$

For case 1. let $r^* := r[z := (\lambda xp)q]$ and $r' := r[z := p[x := q]]$. We prove 1. by induction on r^*. W.l.o.g. assume $z \notin \mathrm{fvar}(p, q) \cup \{x\}$.

i) $\mathrm{rr}(r) = (\lambda xs)t$. By Lemma 2.4 we know

$$\mathrm{redex}(r^*) = \mathrm{redex}(r)^* \text{ and } \mathrm{coat}_{r^*}(\star) = \mathrm{coat}_r(\star)^* \tag{2}$$

thus $\mathrm{rr}(r^*) = \mathrm{rr}(r)^* = (\lambda xs^*)t^*$. Analogously for r'. Hence

$$\begin{aligned}\#r^* &= \mathrm{coat}_{r^*}(s^*[x := t^*]) + \#t^* + 1\\ &\overset{(2)}{=} \mathrm{coat}_r(s[x := t])^* + \#t^* + 1 \overset{(i.h.)}{>} \mathrm{coat}_r(s[x := t])' + \#t' + 1 \overset{sim.}{=} r'\end{aligned}$$

Observe that the induction hypothesis is applied at least once because $z \in \mathrm{fvar}(\mathrm{coat}_r(s[x := t]), t)$.

ii) $\mathrm{rr}(r) = \mathrm{R}\,0ab$, $\mathrm{rr}(r) = \mathrm{R}(\mathrm{S}\,t)ab$ and $\mathrm{rr}(r) = y \in \mathcal{V} \cup \{0, \mathrm{S}, \mathrm{R}\}$ for $y \neq z$. The proofs are the same as in i), because in these cases we also have (2).

iii) $\mathrm{rr}(r) = z$. If $\mathrm{redex}(r) = z\vec{t}$ then $r = z\vec{t}$ by Lemma 2.5. By assumption $z \notin \mathrm{fvar}(\vec{t})$, hence

$$\#r^* = \#(\lambda xp)q\vec{t} = \#p[x := q]\vec{t} + \#q + 1 > \#p[x := q]\vec{t} = \#r'.$$

The other case is $\mathrm{redex}(r) = \mathrm{R}(z\vec{t})ab$. Then we obtain by Lemma 2.5

$$\mathrm{redex}(r^*) = \mathrm{redex}(\mathrm{R}(z\vec{t})^*ab)^* \text{ and } \mathrm{coat}_{r^*}(\star) = \mathrm{coat}_r(\mathrm{coat}_{\mathrm{R}(z\vec{t})^*ab}(\star))^* \tag{3}$$

Again $z \notin \mathrm{fvar}(\mathrm{coat}_r, \vec{t}, a, b)$, thus we compute

$$\begin{aligned}\mathrm{redex}(r^*) &= \mathrm{redex}(\mathrm{R}((\lambda xp)q\vec{t})ab) = (\lambda xp)q\\ \mathrm{coat}_{r^*}(\star) &= \mathrm{coat}_r(\mathrm{R}(\star\vec{t})ab),\end{aligned}$$

hence

$$\begin{aligned}\#r^* &= \#\,\mathrm{coat}_{r^*}(p[x := q]) + \#q + 1\\ &> \#\,\mathrm{coat}_r(\mathrm{R}(p[x := q]\vec{t})ab) = \#\,\mathrm{coat}_r(\mathrm{redex}(r))' = \#r'.\end{aligned}$$

This proves 1. The cases 3. and 4. are proven the same way.

For 2. let $r^* := r[z := \lambda x.px]$ and $r' := r[z := p]$. Again the proof is by induction on r^*. W.l.o.g. assume $z \notin \mathrm{fvar}(p) \cup \{x\}$. If $\mathrm{rr}(r) \neq z$ we proceed as in the proof of 1.

Assume $\mathrm{rr}(r) = z$. If $\mathrm{redex}(r) = z\vec{t}$ then $r = z\vec{t}$ by Lemma 2.5 and $z \notin \mathrm{fvar}(\vec{t})$. First assume $r = z$. Then

$$\#r^* = \#\lambda x.px > \#px \overset{*5}{\geq} \#p = \#r'.$$

At $*5$ we used Lemma 2.7.

Otherwise $r = zt_0\vec{t}$. Hence

$$\#r^* = \#(\lambda x.px)t_0\vec{t} = \#pt_0\vec{t} + \#t_0 + 1 > \#pt_0\vec{t} = \#r'.$$

If $\mathrm{redex}(r) \neq z\vec{t}$ then $\mathrm{redex}(r) = \mathrm{R}(z\vec{t})ab$ by Lemma 2.5. As $\mathrm{lev}(z\vec{t}) = 0$ we must have $\vec{t} = u_0\vec{u}$. Again we obtain by Lemma 2.5 the equations (3), thus we compute $\mathrm{redex}(r^*) = \mathrm{redex}(\mathrm{R}((\lambda x.px)u_0\vec{u})ab) = (\lambda x.px)u_0$ and $\mathrm{coat}_{r^*}(\star) = \mathrm{coat}_r(\mathrm{R}(\star\vec{u})ab)$, hence

$$\#r^* = \#\,\mathrm{coat}_{r^*}(pu_0) + \#u_0 + 1 > \#\,\mathrm{coat}_r(\mathrm{R}(pu_0\vec{u})ab) = \#r'.$$

This proves 2. QED.

Estimate Lemma 2.9 $\overset{\alpha}{\models} t \Rightarrow \#t \leq 2^\alpha$

Proof. We prove by induction on the definition of $\overset{\alpha}{\models} t$

$$\overset{\alpha}{\models} t \Rightarrow \#t \leq 2^\alpha - 1.$$

i) $\mathrm{rr}(t) \in \mathcal{V} \cup \{0, \mathrm{S}, \mathrm{R}\}$. Let $n := \#\,\mathrm{mat}(t)$, then there is a β such that $\beta + n \leq \alpha$ and $\forall s \in \mathrm{mat}(t) \overset{\beta}{\models} s$. We compute

$$\#t = \sum_{s \in \mathrm{mat}(t)} \#s \overset{i.h.}{\leq} \sum_{s \in \mathrm{mat}(t)} (2^\beta - 1) =: m$$

If $n = 0$ then $m = 0 \leq 2^\alpha - 1$. Otherwise

$$m \leq n \cdot (2^\beta - 1) \leq n \cdot 2^\beta - 1 \leq 2^{\beta+n} - 1 \leq 2^\alpha - 1.$$

ii) $\mathrm{rr}(t) = (\lambda xr)s$. There is some $\beta < \alpha$ such that $\overset{\beta}{\models} \mathrm{coat}_t(r[x := s])$ and $\overset{\beta}{\models} s$. Hence

$$\begin{aligned}\#t &= \#\,\mathrm{coat}_t(r[x := s]) + \#s + 1 \overset{i.h.}{\leq} (2^\beta - 1) + (2^\beta - 1) + 1\\ &= 2^{\beta+1} - 1 \leq 2^\alpha - 1.\end{aligned}$$

iii) $\mathrm{rr}(t) = \mathrm{R}\,0ab$. There is some $\beta < \alpha$ such that $\overset{\beta}{\models} \mathrm{coat}_t(b)$ and $\overset{\beta}{\models} a$. Hence

$$\#t = \#\,\mathrm{coat}_t(b) + \#a + 1 \overset{i.h.}{\leq} (2^\beta - 1) + (2^\beta - 1) + 1 \leq 2^\alpha - 1.$$

iv) $\mathrm{rr}(t) = \mathrm{R}(\mathrm{S}\,s)ab$. There is some $\beta < \alpha$ such that $\overset{\beta}{\models} \mathrm{coat}_t(as(\mathrm{R}\,sab))$. Hence

$$\#t = \#\,\mathrm{coat}_t(as(\mathrm{R}\,sab)) + 1 \overset{i.h.}{\leq} (2^\beta - 1) + 1 \leq 2^\alpha - 1.$$

QED.

Combining the Main Lemma with the Estimate Lemma leads to the desired estimation of derivation lengths.

Estimate Theorem 2.10 $\overset{\alpha}{\models} t \Rightarrow \mathrm{d}(t) \leq 2^\alpha$

Proof. Let $s \in \mathrm{T}(\mathcal{V})$ and $k \in \omega$ such that $t \longrightarrow^k s$. Using the Main Lemma and the Estimate Lemma we obtain $k \leq \#t \leq 2^\alpha$. QED.

3 Formal ordinal terms, deduction relations and hierarchies

In this section we develop in detail the technical machinery that is needed in the proof-theoretical analysis of T in section 4.

Definition 3.1 *Inductive definition of a set of terms $\mathcal{T}$ and a subset $\mathcal{P}$ of $\mathcal{T}$.*

1. $0 \in \mathcal{T}$,
2. $1 \in \mathcal{P}$,
3. $\omega \in \mathcal{P}$,
4. $\alpha_1, \ldots, \alpha_m \in \mathcal{P}\ \&\ m \geq 2 \Rightarrow \langle \alpha_1, \ldots, \alpha_m \rangle \in \mathcal{T}$.
5. $\alpha \in \mathcal{T} \Rightarrow 2^\alpha \in \mathcal{P}$.

For $\alpha \in \mathcal{P}$ we put $\langle \alpha \rangle := \alpha$. Then every $\alpha \in \mathcal{T} \setminus \{0\}$ has the form $\alpha = \langle \alpha_1, \ldots, \alpha_m \rangle$ with $\alpha_1, \ldots, \alpha_m \in \mathcal{P}$ and $m \geq 1$. For $\beta \in \mathcal{T}$ we define $0 + \beta := \beta + 0 := \beta$ and for $0 \neq \alpha = \langle \alpha_1, \ldots, \alpha_m \rangle$ and $0 \neq \beta = \langle \beta_1, \ldots, \beta_n \rangle$ we put $\alpha + \beta := \langle \alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_n \rangle$. We identify 0 with the empty sequence $\langle\,\rangle$. We identify the natural numbers with the elements of $\{0, 1, 1+1, 1+1+1, ...\}$.

Definition 3.2 *Inductive definition of an ordinal $\mathcal{O}(\alpha)$ for $\alpha \in \mathcal{T}$.*

1. $\mathcal{O}(0) := 0$,
2. $\mathcal{O}(1) := 1$,
3. $\mathcal{O}(\omega) := \omega$,
4. $\mathcal{O}(\langle \alpha_1, \ldots, \alpha_m \rangle) := \mathcal{O}(\alpha_1) \# \ldots \# \mathcal{O}(\alpha_m)$.

5. $\mathcal{O}(2^\alpha) := 2^{\mathcal{O}(\alpha)+1}$.

Here the ordinal exponentiation with respect to base 2 is defined as follows. For $\alpha = \omega \cdot \beta + m$ with $m < \omega$ let $2^\alpha := \omega^\beta \cdot 2^m$.

Definition 3.3 *Inductive definition of a deduction relation $\leq_0$ on $\mathcal{T}$.*
$\leq_0$ is the least binary relation on $\mathcal{T}$ such that the following holds (where α is an arbitrary element of $\mathcal{T}$) :

1. $\alpha \leq_0 \alpha + \beta$ *for any* $\beta \in \mathcal{T}$.
2. $\alpha + 1 \leq_0 \alpha + \beta$ *for any* $\beta \in \mathcal{T}$ *such that* $\beta \neq 0$.
3. $\alpha + 2 \leq_0 \alpha + \omega$.
4. $\alpha + 2^\beta + 2^\beta \leq_0 \alpha + 2^{\beta+1}$.
5. $\alpha + \beta + 1 \leq_0 \alpha + 1 + \beta$.
6. *If* $\beta \leq_0 \gamma$ *then* $\beta + \delta \leq_0 \gamma + \delta$
7. *If* $\beta \leq_0 \gamma$ *then* $\alpha + 2^\beta \leq_0 \alpha + 2^\gamma$.

Lemma 3.4 1. $\alpha \leq_0 \beta \Rightarrow \gamma + \alpha \leq_0 \gamma + \beta$.

2. $\alpha + k + \beta + l \leq_0 k + l + \alpha + \beta$.
3. $\alpha \leq_0 1 + \alpha$.

Definition 3.5 1. *Let* $N0 := 0$ *and* $N\alpha := n + N\alpha_1 + \cdots + N\alpha_m$ *if* $\varepsilon_0 > \alpha = \omega^{\alpha_1} + \cdots + \omega^{\alpha_m} > \alpha_1 \geq \ldots \geq \alpha_m$.

2. *Let* $F_0(x) := 2^x$ *and* $F_{n+1}(x) := F_n^{x+1}(x)$.
3. *Let* $\Psi(0) := 0$ *and for nonzero* β *let* $\Psi(\beta) := \max\{\Psi(\gamma)+1 : \gamma < \beta \;\&\; N\gamma \leq \Phi(N(\beta))\}$ *where* $\Phi(x) := F_3(x + 3)$.

Lemma 3.6 1. $\alpha < \beta \;\&\; N(\alpha) \leq \Phi(N(\beta)) \Rightarrow \Psi\alpha < \Psi\beta$.

2. $\Psi(\alpha\#\Psi(\beta)) \leq \Psi(\alpha\#\beta)$.
3. $\Psi(k) = k$.
4. $\alpha \geq \omega \Rightarrow \Psi(\alpha) \geq \Phi(N\alpha)$.

Proof. Only assertion 2) needs a proof. The proof of 2) proceeds via induction on β. Assume without loss of generality that $\alpha \neq 0 \neq \beta$. Then

$$\Psi(\alpha\#\Psi(\beta)) = \Psi(\alpha\#\Psi(\gamma) + 1)$$

for some $\gamma < \beta$ such that $N(\gamma) \leq \Phi(N(\beta))$. The induction hypothesis yields

$$\Psi(\alpha\#\Psi(\gamma) + 1) = \Psi(\alpha\#1 + \Psi(\gamma)) \leq \Psi(\alpha\#1\#\gamma).$$

If $\gamma + 1 = \beta$ then we are done. Otherwise $\gamma + 1 < \beta$, hence $\alpha\#1\#\gamma < \alpha\#\beta$ and

$$N(\alpha\#1\#\gamma) = N(\alpha)\#N(1\#\gamma) \leq N(\alpha) + 1 + \Phi(N(\beta)) < \Phi(N(\alpha\#\beta)).$$

Thus assertion 1) yields $\Psi(\alpha\#1\#\gamma) < \Psi(\alpha\#\beta)$. QED.

The function $k \mapsto \psi(\alpha + k)$ is α-descent recursive as can be seen from [BCW94]. More directly this follows from the next lemma.

Definition 3.7 *Let λ be a limit ordinal.*

$$\lambda[k] := \max\{\alpha < \lambda : N(\alpha) \leq \Phi(N\lambda + k)\}$$

Lemma 3.8 *Let λ be a limit ordinal. Then $\Psi(\lambda + k) = \Psi(\lambda[k]) + 1$*

Proof. We have $N(\lambda[k] = \Phi(N\lambda + k)$ since λ is a limit. Thus $\Psi(\lambda + k) \geq \Psi(\lambda[k] + 1)$. We show $\Psi(\lambda + k) \leq \Psi(\lambda[k] + 1)$ by induction on k. Assume $\Psi(\lambda + k) = \Psi(\alpha) + 1$ with $\alpha < \lambda + k$ and $N(\alpha) \leq \Phi(N(\lambda) + k)$. If $\alpha = \lambda + m$ with $m < k$ then the induction hypothesis yields $\Psi(\alpha) \leq \Psi(\lambda[m]) < \Psi(\lambda[k])$ since $\lambda[m] < \lambda[k]$ and $N(\lambda[m]) \leq \Phi(N(\lambda[k]))$. Thus $\Psi(\lambda+k) \leq \Psi(\lambda[k])$. Assume now $\alpha < \lambda$. Then $\alpha \leq \lambda[k]$ by the definition of $\lambda[k]$ and $N(\alpha) \leq \Phi(N(\lambda[k]))$. Hence $\Psi(\alpha) \leq \Psi(\lambda[k])$. QED.

Definition 3.9 *Recursive definition of a natural number $\mathcal{D}(\alpha)$ for $\alpha \in \mathcal{T}$.*

1. $\mathcal{D}(0) := 0$,
2. $\mathcal{D}(1) := 1$,
3. $\mathcal{D}(\omega) := \Psi(\omega)$,
4. $\mathcal{D}(2^\alpha) := \Psi(2^{\mathcal{O}(\alpha)+1})$.
5. $\mathcal{D}(\langle\alpha_1, \dots, \alpha_m\rangle) :=$
   $\Psi(\mathcal{O}(\alpha_m) + \Psi(\mathcal{O}(\alpha_{m-1}) + \Psi(... + \Psi(\mathcal{O}(\alpha_2) + \Psi(\mathcal{O}(\alpha_1)))...)))$.

Then we have $\mathcal{D}(\langle\alpha_1, \dots, \alpha_m\rangle) = \Psi(\mathcal{O}(\alpha_m) + \mathcal{D}(\langle\alpha_1, \dots, \alpha_{m-1}\rangle)$ and $\mathcal{D}(\alpha+1) = \mathcal{D}(\alpha) + 1$.

Lemma 3.10

1. $N(2^\alpha) \leq 2^{N\alpha}$,
2. $N(\alpha) + 1 \leq N(2^{\alpha+1}), N(\alpha) \leq N(2^\alpha) \cdot 2$,
3. $\alpha \leq_0 \beta \Rightarrow N(\mathcal{O}(\alpha)) \leq F_2(N(\mathcal{O}(\beta)))$.

Proof. Assertions 1) and 2) are easy to prove. Assertion 3) follows by an induction along the inductive definition of $\leq_0$. For the critical case assume that $\alpha = \gamma + 2^{\alpha'}$, $\beta = \gamma + 2^{\beta'}$ and $\alpha' \leq_0 \beta'$. Then the induction hypothesis yields

$$\begin{aligned} N(\mathcal{O}(\alpha)) \leq N(\mathcal{O}(\gamma) + 2^{N(\mathcal{O}(\alpha'))+1}) &\leq N(\mathcal{O}(\gamma) + 2^{F_2(N(\mathcal{O}(\beta')))+1}) \\ &< N(\mathcal{O}(\gamma)) + F_2(N(\mathcal{O}(\beta')) + 1) \\ &\leq N(\mathcal{O}(\gamma)) + F_2(N(\mathcal{O}(2^{\beta'})) \\ &\leq F_2(N(\mathcal{O}(\gamma + 2^{\beta'}))) \end{aligned}$$

QED.

Lemma 3.11 $\alpha \leq_0 \beta \quad \Longrightarrow \quad \mathcal{D}(\alpha) \leq \mathcal{D}(\beta)$.

Proof by an induction along the inductive definition of $\leq_0$.
1. Assume that

$$\alpha \leq_0 \beta = \alpha + \langle \gamma_1, \dots, \gamma_m \rangle$$

where $m \geq 0$ and $\gamma_1, \dots, \gamma_m \in \mathcal{P}$. Then

$$\begin{array}{rcl} \mathcal{D}(\beta) & = & \mathcal{D}(\alpha + \langle \gamma_1, \dots, \gamma_m \rangle) \\ & = & \Psi(\mathcal{O}(\gamma_m) + \mathcal{D}(\alpha + \langle \gamma_1, \dots, \gamma_{m-1} \rangle)) \\ & \geq & \mathcal{D}(\alpha + \langle \gamma_1, \dots, \gamma_{m-1} \rangle) \\ & \geq & \dots \\ & \geq & \mathcal{D}(\alpha + \gamma_1) \\ & = & \Psi(\mathcal{O}(\gamma_1) + \mathcal{D}(\alpha)) \\ & \geq & \mathcal{D}(\alpha). \end{array}$$

2. Assume that $\alpha = \alpha' + 1$ and $\beta = \alpha' + \langle \gamma_1, \dots, \gamma_m \rangle$ where $m \geq 1$ and $\gamma_1, \dots, \gamma_m \in \mathcal{P}$. Then

$$\begin{array}{rcl} \mathcal{D}(\beta) & = & \mathcal{D}(\alpha' + \langle \gamma_1, \dots, \gamma_m \rangle) \\ & = & \Psi(\mathcal{O}(\gamma_m) + \mathcal{D}(\alpha' + \langle \gamma_1, \dots, \gamma_{m-1} \rangle) \\ & \geq & \mathcal{D}(\alpha' + \langle \gamma_1, \dots, \gamma_{m-1} \rangle) \\ & \geq & \dots \\ & \geq & \mathcal{D}(\alpha' + \gamma_1) \\ & = & \Psi(\mathcal{O}(\gamma_1) + \mathcal{D}(\alpha')) \\ & \geq & 1 + \mathcal{D}(\alpha') = \mathcal{D}(\alpha' + 1) = \mathcal{D}(\alpha). \end{array}$$

3. Assume that $\alpha = \alpha' + 2$ and $\beta = \alpha' + \omega$. Then

$$\mathcal{D}(\beta) = \mathcal{D}(\alpha' + \omega) = \Psi(\omega + \mathcal{D}(\alpha')) > 2 + \mathcal{D}(\alpha') = \mathcal{D}(\alpha).$$

4. Assume that $\alpha = \alpha' + 2^\gamma + 2^\gamma$ and $\beta = \alpha' + 2^{\gamma+1}$. Then

$$\begin{array}{rcl} \mathcal{D}(\beta) & = & \mathcal{D}(\alpha' + 2^{\gamma+1}) \\ & = & \Psi(2^{\mathcal{O}(\gamma)+1+1} + \mathcal{D}(\alpha')) \\ & = & \Psi(2^{\mathcal{O}(\gamma)+1} \# 2^{\mathcal{O}(\gamma)+1} + \mathcal{D}(\alpha')) \\ & \geq & \Psi(2^{\mathcal{O}(\gamma)+1} \# \Psi(2^{\mathcal{O}(\gamma)+1} + \mathcal{D}(\alpha'))) \\ & = & \mathcal{D}(\alpha' + 2^\gamma + 2^\gamma) = \mathcal{D}(\alpha). \end{array}$$

5. Assume that $\alpha = \alpha' + \langle \beta_1, \dots, \beta_m \rangle + 1$ and $\beta = \alpha' + 1 + \langle \beta_1, \dots, \beta_m \rangle$. Then

$$\begin{array}{rcl} \mathcal{D}(\beta) & = & \Psi(\mathcal{O}(\beta_1) + \Psi(... + \Psi(\mathcal{O}(\beta_m) + \mathcal{D}(\alpha' + 1))...)) \\ & \geq & \Psi(\mathcal{O}(\beta_1) + \Psi(... + \Psi(\mathcal{O}(\beta_m) + 1 + \mathcal{D}(\alpha')...)) \\ & \geq & \Psi(\mathcal{O}(\beta_1) + \Psi(...\Psi(\mathcal{O}(\beta_{m-1}) + 1 + \Psi(\mathcal{O}(\beta_m) + \mathcal{D}(\alpha')))...)) \\ & \geq & \Psi(\mathcal{O}(\beta_1) + 1 + \Psi(...\Psi(\mathcal{O}(\beta_{m-1}) + \Psi(\mathcal{O}(\beta_m) + \mathcal{D}(\alpha')))...)) \\ & \geq & 1 + \Psi(\mathcal{O}(\beta_1) + \Psi(...\Psi(\mathcal{O}(\beta_{m-1}) + \Psi(\mathcal{O}(\beta_m) + \mathcal{D}(\alpha')))...) \\ & = & \mathcal{D}(\alpha). \end{array}$$

6. Assume that $\alpha = \alpha' + \delta$, $\beta = \beta' + \delta$ where $\alpha' \leq_0 \beta'$ and $\delta = \langle \delta_1, \dots, \delta_n \rangle$ with $n \geq 0$ and $\delta_1, \dots, \delta_n \in \mathcal{P}$. The induction hypothesis yields $\mathcal{D}(\alpha') \leq \mathcal{D}(\beta')$.

Then

$$\begin{aligned}
\mathcal{D}(\alpha) &= \psi(\mathcal{O}(\delta_n) + \psi(... + \psi(\mathcal{O}(\delta_1) + \mathcal{D}(\alpha'))...)) \\
&\geq \psi(\mathcal{O}(\delta_n) + \psi(... + \psi(\mathcal{O}(\delta_1) + \mathcal{D}(\beta'))...)) \\
&= \mathcal{D}(\beta).
\end{aligned}$$

7. Assume now that $\alpha = \gamma + 2^{\alpha'}$, $\beta = \gamma + 2^{\beta'}$ and $\alpha' \leq_0 \beta'$. Then $\mathcal{O}(\alpha') \leq \mathcal{O}(\beta')$. If $\mathcal{O}(\alpha') = \mathcal{O}(\beta')$ then $\mathcal{D}(\alpha) = \mathcal{D}(\beta)$. We may thus assume that $\mathcal{O}(\alpha') < \mathcal{O}(\beta')$. Then

$$2^{\mathcal{O}(\alpha')+1} + \mathcal{D}(\gamma) < 2^{\mathcal{O}(\beta')+1} + \mathcal{D}(\gamma).$$

The assumption $\alpha' \leq_0 \beta'$ yields $N(\mathcal{O}(\alpha')) \leq F_2(N(\mathcal{O}(\beta')))$ hence

$$\begin{aligned}
N(2^{\mathcal{O}(\alpha')+1} + \mathcal{D}(\gamma)) &\leq 2^{F_2(N(\mathcal{O}(\beta')))+1} + \mathcal{D}(\gamma) \\
&\leq F_2(N(2^{\mathcal{O}(\beta')+1})) + \mathcal{D}(\gamma) \\
&\leq \Phi(N(2^{\mathcal{O}(\beta')+1} + \mathcal{D}(\gamma)))
\end{aligned}$$

Therefore assertion 1) of Lemma 3.6 yields

$$\mathcal{D}(\alpha) = \mathcal{D}(\gamma + 2^{\alpha'}) < \mathcal{D}(\gamma + 2^{\beta'}) = \mathcal{D}(\beta).$$

QED.

Definition 3.12 *Inductive definition of a set $\mathcal{C}$ of contexts.*

1. $\alpha + \star \in \mathcal{C}$ *for any* $\alpha \in \mathcal{T}$.
2. $f \in \mathcal{C} \Rightarrow \alpha + 2^f \in \mathcal{C}$ *for any* $\alpha \in \mathcal{T}$.

For $\alpha \in \mathcal{T}$ we denote by $f(\alpha)$ the result of substituting the placeholder $\star$ in f by α. The result $f(\alpha)$ is then an element of $\mathcal{T}$.

Lemma 3.13 *Assume that* $f \in \mathcal{C}$.

1. $\mathcal{O}(f(k)) + l < \mathcal{O}(f(\omega))$ *for any* $k, l < \omega$.
2. $N(\mathcal{O}(f(k))) \leq F_2(N(\mathcal{O}(f(\omega))) + k)$.
3. $\Psi(\alpha \# 2^{\mathcal{O}(f(k))+1} + l) < \Psi(\alpha \# 2^{\mathcal{O}(f(\omega))} + k)$ *for any* $k, l < \omega$ *such that* $l \leq k$.

Proof. 1. Assume first that $f = \alpha + \star$. Then

$$\mathcal{O}(f(k)) + l = \mathcal{O}(\alpha) \# k + l < \mathcal{O}(\alpha) \# \omega.$$

Assume now that $f = \alpha + 2^g$. Then the induction hypothesis yields

$$\begin{aligned}
\mathcal{O}(f(k)) + l &= \mathcal{O}(\alpha) \# 2^{\mathcal{O}(g(k))+1} + l \\
&\leq \mathcal{O}(\alpha) \# 2^{\mathcal{O}(g(k))+1+l} \\
&< \mathcal{O}(\alpha) \# 2^{\mathcal{O}(g(\omega))+1} \\
&= \mathcal{O}(\alpha + 2^{g(\omega)}).
\end{aligned}$$

2. Assume first that $f = \alpha + \star$. Then

$$\begin{array}{rcl} N(\mathcal{O}(f(k))) & = & N(\mathcal{O}(\alpha) + k) \\ & < & F_2(N(\mathcal{O}(\alpha)\#\omega)) + k) \\ & = & F_2(N(\mathcal{O}(f(\omega))) + k). \end{array}$$

Assume now that $f = \alpha + 2^g$. Then the induction hypothesis yields

$$\begin{array}{rcl} N(\mathcal{O}(f(k))) & = & N(\mathcal{O}(\alpha) + N(2^{\mathcal{O}(g(k))+1}) \\ & \leq & N(\mathcal{O}(\alpha)) + 2^{F_2(N(\mathcal{O}(g(\omega)))+k)+1} \\ & \leq & N(\mathcal{O}(\alpha)) + F_2(N(\mathcal{O}(g(\omega))) + k + 1) \\ & \leq & N(\mathcal{O}(\alpha)) + F_2(N(2^{\mathcal{O}(g(\omega))+1} + k) \\ & \leq & F_2(N(\mathcal{O}(\alpha + 2^{g(\omega)}))) + k) \end{array}$$

3. Assertion 1) yields

$$\alpha\#2^{\mathcal{O}(f(k))+1} + l < \alpha\#2^{\mathcal{O}(f(k))+1+l} < \alpha\#2^{\mathcal{O}(f(\omega))}.$$

Assertion 2) yields

$$\begin{array}{rcl} N(\alpha\#2^{\mathcal{O}(f(k))+1} + l) & \leq & N(\alpha)\#2^{N(\mathcal{O}(f(k)))+1} + l \\ & \leq & N(\alpha)\#2^{F_2(N(\mathcal{O}(f(\omega))+k))+1} + l \\ & \leq & N(\alpha)\#2^{F_2(N(2^{\mathcal{O}(f(\omega))})\cdot 2+k))+1} + l \\ & \leq & \Phi(N(\alpha\#2^{\mathcal{O}(f(\omega))} + k)) \end{array}$$

since $N\alpha \leq N(2^\alpha) \cdot 2$.

The assertion follows by assertion 1) of Lemma 3.6 QED.

Lemma 3.14 $f \in \mathcal{C} \Rightarrow \mathcal{D}(f(\mathcal{D}f(0))) < \mathcal{D}(f(\omega))$.

Proof. Assume first that $f = \alpha + \star$. Then

$$\mathcal{D}(f(\mathcal{D}(f(0)))) = \mathcal{D}(\alpha + \mathcal{D}(\alpha)) = \mathcal{D}(\alpha) \cdot 2$$

and

$$\begin{array}{l} \mathcal{D}(f(\omega)) = \mathcal{D}(\alpha + \omega) = \psi(\omega + \mathcal{D}(\alpha)) \\ \qquad \geq \Phi(N(\omega + \mathcal{D}(\alpha))) > \mathcal{D}(\alpha) \cdot 2 = \mathcal{D}(f(\mathcal{D}(f(0)))). \end{array}$$

Assume now that $f = \alpha + 2^g$. Then assertion 2) of Lemma 3.6 and assertion 3) of Lemma 3.13 yield

$$\begin{array}{rcl} \mathcal{D}(f(\mathcal{D}(f(0)))) & = & \mathcal{D}(\alpha + 2^{g(\mathcal{D}(\alpha+2^{g(0)})))}) \\ & = & \Psi(2^{\mathcal{O}(g(\Psi(2^{\mathcal{O}(g(0))+1}+\mathcal{D}(\alpha)))))+1} + \mathcal{D}(\alpha)) \\ & < & \Psi(2^{\mathcal{O}(g(\omega)} + \Psi(2^{\mathcal{O}(g(0))+1} + \mathcal{D}(\alpha))) \\ & \leq & \Psi(2^{\mathcal{O}(g(\omega)} + 2^{\mathcal{O}(g(0))+1} + \mathcal{D}(\alpha)) \\ & < & \Psi(2^{\mathcal{O}(g(\omega)} + 2^{\mathcal{O}(g(\omega))} + \mathcal{D}(\alpha)) \\ & = & \Psi(2^{\mathcal{O}(g(\omega)+1} + \mathcal{D}(\alpha)) = \mathcal{D}(f(\omega)). \end{array}$$

QED.

4 Adding cut-rule and BUCHHOLZ' Ω-rule

Our strategy for estimating $\mathrm{d}(r)$ is to compute the expanded head reduction tree of r. Therefore we extend the expanded head reduction trees by a cut-rule and an appropriate miniaturization of BUCHHOLZ' Ω-rule which allow a simple embedding of any term of GÖDEL's T into the extended calculus. Then we first eliminate cuts and afterwards the Ω-rule by adapting collapsing techniques from BUCHHOLZ' treatment of ID_1 (cf. [Bu80]). In this way we obtain expanded head reduction trees for any term of GÖDEL's T with an optimal upper bound on its size.

The above mentioned Ω-rule will have the following form: If

$$\forall k \in \omega \forall t \in \mathrm{T}(\mathcal{V})\Big(\overset{k}{\models} t \text{ and } t \text{ of type } 0 \;\Rightarrow\; \overset{f[k]}{\vdash} \mathrm{R}\, tab\Big)$$

then $\overset{f[\omega]}{\vdash} \mathrm{R}$ where a, b are suitable variables. We should observe at this point the special meaning of the terms t in this context. They are in some sense bounded, especially the variables which occur in such a term serve rather as a parameter than a variable. This means that during the cut-elimination procedure, where cuts are replaced by substitutions, these parameter-variables are not allowed to be substituted because the Ω-rule is not robust under such substitution. From this it follows that also recursors which occur in such terms have another meaning than those which are derived via Ω-rule, i.e. they can be derived as before. In order to model this difference technically we need a copy $\mathrm{T}'(\mathcal{V})$ of $\mathrm{T}(\mathcal{V})$ for which substitution can be handled.

Let $\mathcal{V}' := \{v' : v \in \mathcal{V}\}$ be a distinct copy of $\mathcal{V}$. Let $\overline{\mathcal{V}} := \mathcal{V} \cup \mathcal{V}'$ and define $\mathcal{V}'_{\mathrm{R}}, \mathcal{V}'_{0,\mathrm{S}}, \mathcal{V}'_{0,\mathrm{S},\mathrm{R}}, \overline{\mathcal{V}}_{\mathrm{R}}, \overline{\mathcal{V}}_{0,\mathrm{S}}, \overline{\mathcal{V}}_{0,\mathrm{S},\mathrm{R}}$ analogously to $\mathcal{V}_{\mathrm{R}}, \mathcal{V}_{0,\mathrm{S}}, \mathcal{V}_{0,\mathrm{S},\mathrm{R}}$. Let R' be a new symbol and define $\overline{\mathrm{R}} := \{\mathrm{R}, \mathrm{R}'\}$. Observe that $\mathrm{R}' \notin \mathcal{V}'_{\mathrm{R}}, \overline{\mathcal{V}}_{\mathrm{R}}$ etc.

With $\overline{x}$ we mean x or x' for $x \in \mathcal{V}$.

A ground type ι has *level* $\mathrm{lev}(\iota) = 0$ and $\mathrm{lev}(\rho \to \sigma) = \max(\mathrm{lev}(\rho)+1, \mathrm{lev}(\sigma))$. The *level* $\mathrm{lev}(r)$ of r is defined to be the level $\mathrm{lev}(\sigma)$ of its type σ, the *degree* $\mathrm{g}(r)$ of r is defined to be the maximum of the levels of subterms of r.

Definition 4.1 *We define* $\mathrm{T}'(\mathcal{V})$ *inductively by*

- $\mathcal{V}' \cup \{0, \mathrm{S}, \mathrm{R}'\} \subset \mathrm{T}'(\mathcal{V})$
- $r, s \in \mathrm{T}'(\mathcal{V})$ *and* $x \in \mathcal{V}' \Rightarrow (\lambda x r), (rs) \in \mathrm{T}'(\mathcal{V})$
- $t \in \mathrm{T}(\mathcal{V})$ *and* $\mathrm{lev}(t) = 0 \Rightarrow (\mathrm{R}'\, t) \in \mathrm{T}'(\mathcal{V})$

Let $\overline{\mathrm{T}}(\mathcal{V}) := \mathrm{T}'(\mathcal{V}) \cup \mathrm{T}(\mathcal{V})$.

There are two canonical mappings, the embedding $\overline{\cdot} : \mathrm{T}(\mathcal{V}) \to \overline{\mathrm{T}}(\mathcal{V})$ *and the breakup* $\widehat{\cdot} : \overline{\mathrm{T}}(\mathcal{V}) \to \mathrm{T}(\mathcal{V})$ *which are recursively defined by*

- $\overline{x} := x'$ *and* $\widehat{x'} := x$ *for* $x \in \mathcal{V}_{\mathrm{R}}$
- $\overline{0} := 0$, $\overline{\mathrm{S}} := \mathrm{S}$ *and* $\widehat{x} := x$ *for* $x \in \mathcal{V}_{0,\mathrm{S}}$
- $\overline{\lambda x r} := \lambda x' \overline{r}$ *for* $x \in \mathcal{V}$, $\overline{rs} := \overline{r}\,\overline{s}$.
- $\widehat{\lambda \overline{x} r} := \lambda x \widehat{r}$ *for* $x \in \mathcal{V}$, $\widehat{rs} := \widehat{r}\,\widehat{s}$.

Obviously $\widehat{\overline{t}} = t$ *for* $t \in \mathrm{T}(\mathcal{V})$.

We are considering λ-terms only modulo α-conversion without making this too explicit. Of course sometimes this causes problems, e.g. in defining $\widehat{\lambda\overline{x}r} := \lambda x\widehat{r}$ we have to make sure x does not occur in r. One way obtaining this is to define $\widehat{\lambda x' r} := \lambda y.(r[x' := y'])\widehat{}$ for some $y \in \mathcal{V}$ such that y, y' do not occur in r. Another possibility is – and we will consider this in the following – to assume always x not occurring in r when writing $\lambda x' r$ for $x \in \mathcal{V}$.

We state some simple observations about the relationship of $\mathrm{T}(\mathcal{V})$, $\mathrm{T}'(\mathcal{V})$ and $\overline{\mathrm{T}}(\mathcal{V})$.

1. $\mathrm{T}'(\mathcal{V}) \cap \mathrm{T}(\mathcal{V}) = \{\mathrm{S}\} \cup \{\mathrm{S}^k\, 0 \;:\; k \in \omega\}$
2. $(rs) \in \mathrm{T}'(\mathcal{V}) \Rightarrow r \in \mathrm{T}'(\mathcal{V})$
3. $\overline{\mathrm{T}}(\mathcal{V})$ is closed under subterms.
4. $(rs) \in \overline{\mathrm{T}}(\mathcal{V})$ and $r \in \mathrm{T}(\mathcal{V}) \Rightarrow r = \mathrm{S}$ or $s \in \mathrm{T}(\mathcal{V})$
   Proof. If $(rs) \in \mathrm{T}'(\mathcal{V})$ then $r \in \mathrm{T}'(\mathcal{V}) \cap \mathrm{T}(\mathcal{V})$ thus $r = \mathrm{S}$. Otherwise $(rs) \in \mathrm{T}(\mathcal{V})$, hence $s \in \mathrm{T}(\mathcal{V})$. QED.
5. $(rs) \in \overline{\mathrm{T}}(\mathcal{V})$ and $s \in \mathrm{T}(\mathcal{V}) \setminus \mathrm{T}'(\mathcal{V}) \Rightarrow r \in \mathrm{T}(\mathcal{V})$ or $r = \mathrm{R}'$
6. $r \in \overline{\mathrm{T}}(\mathcal{V})$, $x \in \mathcal{V}'$ and $s \in \mathrm{T}'(\mathcal{V}) \Rightarrow r[x := s] \in \overline{\mathrm{T}}(\mathcal{V})$ and
   $r \in \overline{\mathrm{T}}(\mathcal{V})$, $x \in \mathcal{V}$ and $s \in \mathrm{T}(\mathcal{V}) \Rightarrow r[x := s] \in \overline{\mathrm{T}}(\mathcal{V})$
7. $\mathrm{redex}(t), s \in \mathrm{T}(\mathcal{V}) \Rightarrow \mathrm{coat}_t(s) \in \overline{\mathrm{T}}(\mathcal{V})$ and
   $\mathrm{redex}(t), s \in \mathrm{T}'(\mathcal{V}) \Rightarrow \mathrm{coat}_t(s) \in \overline{\mathrm{T}}(\mathcal{V})$

Definition 4.2 *We extend the definition of* $\mathrm{redex}(s), \mathrm{rr}(s), \mathrm{coat}_s(\star)$ *and* $\mathrm{mat}(s)$ *to* $s \in \overline{\mathrm{T}}(\mathcal{V})$.

$s \in \mathrm{T}'(\mathcal{V})$		$\mathrm{rr}(s)$	$\mathrm{coat}_s(\star)$	$\mathrm{mat}(s)$
$x\vec{t}$	$x \in \mathcal{V}'_{0,\mathrm{S}}$	x	$\star$	$\{\{\vec{t}\}\}$
λxr		$(\lambda xr)x$	$\star$	$\emptyset$
$(\lambda xr)u\vec{t}$		$(\lambda xr)u$	$\star\vec{t}$	$\emptyset$
$\mathrm{R}'\, u_1 \dots u_l$	$l \leq 2$	R'	$\star$	$\{\{\vec{u}\}\}$
$\mathrm{R}'\, tab\vec{s}$:				
$t = 0, \mathrm{S}\, t'$		$\mathrm{R}'\, tab$	$\star\vec{s}$	$\emptyset$
$t = x\vec{u},\ x \in \overline{\mathcal{V}}$		R'	$\star\vec{s}$	$\{\{\vec{u}, a, b, \vec{s}\}\}$
$t \neq x\vec{u},\ x \in \overline{\mathcal{V}}_{0,\mathrm{S}}$		$\mathrm{rr}(t)$	$\mathrm{R}'\, \mathrm{coat}_t(\star)ab\vec{s}$	$\mathrm{mat}(t) \cup \{\{a, b, \vec{s}\}\}$

Again we have

$$\mathrm{rr}(t) \in \overline{\mathcal{V}}_{0,\mathrm{S},\mathrm{R}} \cup \{\mathrm{R}'\} \cup \{(\lambda xr)s, \mathrm{R}^*\, 0ab, \mathrm{R}^*(\mathrm{S}\, s)ab \mid a, b, r, s \in \overline{\mathrm{T}}(\mathcal{V}), \mathrm{R}^* \in \overline{\mathrm{R}}\}$$

Furthermore we observe $\mathrm{rr}(\widehat{t}) = \mathrm{rr}(t)\widehat{}$ if $\mathrm{rr}(t) \neq \mathrm{R}'$ and $\mathrm{coat}_{\widehat{t}}(\star) = \mathrm{coat}_t(\star)\widehat{}$.

We now extend $\overset{\alpha}{\models} t$ by cuts and Ω-rules. Let a context $c(\star)$ be a term in which $\star$ occurs exactly once. With $\leq_0^*$ we denote the transitive closure of $\leq_0$.

Definition 4.3 *We inductively define* $\mathop{\vdash}\limits^{\alpha}_{\rho} t$ *for* $t \in \overline{\mathrm{T}}(\mathcal{V})$, $\alpha \in \mathcal{T}$ *and* $\rho < \omega$ *if one of the following cases holds:*

(Acc-Rule) *There is some* β *such that* $\beta \leq_0 \alpha$ *and* $\mathop{\vdash}\limits^{\beta}_{\rho} t$.

($\overline{\mathcal{V}}_{0,\mathrm{S,R}}$-Rule) $\mathrm{rr}(t) \in \overline{\mathcal{V}}_{0,\mathrm{S,R}}$, $\alpha = \beta + \#\,\mathrm{mat}(t)$ *and* $\forall s \in \mathrm{mat}(t) \stackrel{\beta}{\underset{\rho}{|}} s$.

(β-Rule) $\mathrm{rr}(t) = (\lambda xr)s$ $(x \in \overline{\mathcal{V}})$, $\alpha = \beta + 1$ *and* $\stackrel{\beta}{\underset{\rho}{|}} \mathrm{coat}_t(r[x := s])$ *and* $\stackrel{\beta}{\underset{\rho}{|}} s$.

($\overline{\mathrm{R}}0$-Rule) $\mathrm{rr}(t) = \mathrm{R}^*\,0ab$ $(\mathrm{R}^* \in \overline{\mathrm{R}})$, $\alpha = \beta + 1$ *and* $\stackrel{\beta}{\underset{\rho}{|}} \mathrm{coat}_t(b)$ *and* $\stackrel{\beta}{\underset{\rho}{|}} a$.

($\overline{\mathrm{R}}\,\mathrm{S}$-Rule) $\mathrm{rr}(t) = \mathrm{R}^*(\mathrm{S}\,t')ab$ $(\mathrm{R}^* \in \overline{\mathrm{R}})$, $\alpha = \beta + 1$ *and* $\stackrel{\beta}{\underset{\rho}{|}} \mathrm{coat}_t(at'(\mathrm{R}^*\,t'ab))$.

(Cut-Rule) $t = (rs)$, $\mathrm{lev}(r) \leq \rho$, $s \in \mathrm{T}'(\mathcal{V})$, $\alpha = \beta + 1$ *and* $\stackrel{\beta}{\underset{\rho}{|}} r$ *and* $\stackrel{\beta}{\underset{\rho}{|}} s$.

($\mathrm{R}'\,\Omega_0$-Rule) $t = \mathrm{R}'\,u_1 \ldots u_l$, $l \leq 2$, *there are new variables* $u_{l+1}, \ldots, u_3 \in \mathcal{V}'$, *distinct in pairs, and some* $\beta[\star] \in \mathcal{C}$ *such that* $\alpha = \beta[\omega] + 1$, $\beta[0] + 2 \leq_0^* \alpha$, $\stackrel{\beta[0]}{\underset{\rho}{|}} u_i$ *for* $1 \leq i \leq l$ *and*

$$\forall u \in \mathrm{T}(\mathcal{V}) \forall k < \omega \Big(\mathrm{lev}(u) = 0 \;\&\; \stackrel{k}{\models} u \Rightarrow \stackrel{\beta[k]}{\underset{\rho}{|}} \mathrm{R}'\,uu_2u_3 \Big)$$

($\mathrm{R}'\,\Omega_1$-Rule) $t = c(\mathrm{R}'\,sab)$ *for some context* $c(\star)$ *and there is some* $\beta[\star] \in \mathcal{C}$ *such that* $\alpha = \beta[\omega] + 1$, $\stackrel{\beta[0]}{\underset{\rho}{|}} s$ *and*

$$\forall u \in \mathrm{T}(\mathcal{V}) \forall k < \omega \Big(\mathrm{lev}(u) = 0 \;\&\; \stackrel{k}{\models} u \Rightarrow \stackrel{\beta[k]}{\underset{\rho}{|}} c(\mathrm{R}'\,uab) \Big)$$

Structural Rule 4.4 $\stackrel{\alpha}{\underset{\rho}{|}} t, \alpha \leq_0^* \alpha', \rho \leq \rho' \Rightarrow \stackrel{\alpha'}{\underset{\rho'}{|}} t$

Proof. A simple induction on the definition of $\stackrel{\alpha}{\underset{\rho}{|}} t$ shows $\stackrel{\alpha}{\underset{\rho'}{|}} t$, then we apply several times the Acc-Rule. QED.

We observe that the cut-free system is a subsystem of the one with cuts.

Lemma 4.5 $\stackrel{\alpha}{\models} t \Rightarrow \stackrel{\alpha}{\underset{0}{|}} t$

Proof. The proof is a simple induction on the definition of $\stackrel{\alpha}{\models} t$, because $\alpha < \beta < \omega \Rightarrow \alpha + 1 \leq_0^* \beta$. QED.

Variable Substitution Lemma 4.6 *Assume* $\stackrel{\alpha}{\underset{\rho}{|}} t$.

1. $x, y \in \mathcal{V} \Rightarrow \stackrel{\alpha}{\underset{\rho}{|}} t[x := y]$.
2. $x, y \in \mathcal{V}' \Rightarrow \stackrel{\alpha}{\underset{\rho}{|}} t[x := y]$.

For the next lemma observe that $\alpha \leq_0 \alpha + 1 \leq_0 1 + \alpha$ holds for all α.

Appending Lemma 4.7 *Assume* $\stackrel{\alpha}{\underset{\rho}{|}} t$. *If* $y \in \mathcal{V}'$ *and* $ty \in \overline{\mathrm{T}}(\mathcal{V})$ *then* $\stackrel{1+\alpha}{\underset{\rho}{|}} ty$.

Proof. The proof is by induction on the definition of $\stackrel{\alpha}{\underset{\rho}{|}} t$.
Acc-Rule. Follows directly from the induction hypothesis by Acc-Rule and the fact that $\beta \leq_0 \alpha \Rightarrow 1 + \beta \leq_0 1 + \alpha$.

$\overline{\mathcal{V}}_{0,\mathrm{S},\mathrm{R}}$-Rule. $\mathrm{rr}(t) = \mathrm{R}$ is not possible because $t = \mathrm{R}\,u_1 \dots u_l$ would imply $t \in \mathrm{T}(\mathcal{V}) \setminus \mathrm{T}'(\mathcal{V})$ and therefore $ty \notin \overline{\mathrm{T}}(\mathcal{V})$.

In case $\mathrm{rr}(t) \in \overline{\mathcal{V}}_{0,\mathrm{S}}$ the assertion follows because $\mathrm{rr}(ty) = \mathrm{rr}(t)$, $\vdash^{\gamma}_{\rho} y$ for arbitrary γ and $\beta + n = \alpha \Rightarrow \beta + n + 1 = \alpha + 1 \leq_0 1 + \alpha$.

β-Rule. $\mathrm{rr}(t) = (\lambda x r)s$, $\alpha = \beta + 1$ and $\vdash^{\beta}_{\rho} \mathrm{coat}_t(r[x := s])$, $\vdash^{\beta}_{\rho} s$.

If $t = \lambda x r$ then $s = x$, hence $\mathrm{coat}_t(r[x := s]) = r$. Assuming $x \in \mathcal{V}$ would imply $t \in \mathrm{T}(\mathcal{V}) \setminus \mathrm{T}'(\mathcal{V})$ contradicting $ty \in \overline{\mathrm{T}}(\mathcal{V})$, thus $x \in \mathcal{V}'$ and therefore $\vdash^{\beta}_{\rho} r[x := y]$ by the previous Lemma. Hence $\vdash^{\alpha}_{\rho} ty$ with $\vdash^{\beta}_{\rho} y$ and β-Rule, thus $\vdash^{1+\alpha}_{\rho} ty$ with Acc-Rule.

Otherwise $\mathrm{rr}(ty) = \mathrm{rr}(t) = (\lambda x r)s$. We obtain $\vdash^{1+\beta}_{\rho} \mathrm{coat}_t(r[x := s])y$ by induction hypothesis. As $\beta \leq_0 1 + \beta$ we also have $\vdash^{1+\beta}_{\rho} s$. Now $\mathrm{coat}_{ty}(\star) = \mathrm{coat}_t(\star)y$, hence $\vdash^{1+\alpha}_{\rho} ty$ by β-Rule.

$\overline{\mathrm{R}}0$-Rule and $\overline{\mathrm{R}}\,\mathrm{S}$-Rule are similar to β-Rule.

Cut-Rule. $t = rs$ with $\mathrm{lev}(r) \leq \rho$, thus $\mathrm{lev}(t) \leq \rho$, hence $\vdash^{\alpha+1}_{\rho} ty$ by a Cut-Rule and we obtain $\vdash^{1+\alpha}_{\rho} ty$ by a Acc-Rule.

$\mathrm{R}'\,\Omega_0$-Rule. $t = \mathrm{R}'\,u_1 \dots u_l$, $l \leq 2$, there are new variables $u_{l+1}, \dots, u_3 \in \mathcal{V}' \setminus \{y\}$, distinct in pairs, and some $\beta[\star] \in \mathcal{C}$ such that $\alpha = \beta[\omega]+1$, $\beta[0]+2 \leq^*_0 \alpha$, $\vdash^{\beta[0]}_{\rho} u_i$ for $1 \leq i \leq l$ and for $u \in \mathrm{T}(\mathcal{V})$, $k < \omega$ with $\mathrm{lev}(u) = 0$ and $\models^{k} u$ also $\vdash^{\beta[k]}_{\rho} \mathrm{R}'\,uu_2u_3$. Let $u'_1 \dots u'_3$ be $u_1 \dots u_3[u_{l+1} := y]$, then $\vdash^{1+\beta[k]}_{\rho} \mathrm{R}'\,uu'_2u'_3$ by the previous Lemma and Acc-Rule. Let $\gamma[\star] := 1 + \beta[\star]$, then $\gamma[\star] \in \mathcal{C}$, $\gamma[\omega] + 1 = 1 + \beta[\omega] + 1 = 1 + \alpha$ and $\gamma[0] + 2 = 1 + \beta[0] + 2 \leq^*_0 1 + \alpha$, hence $\vdash^{1+\alpha}_{\rho} t$ by $\mathrm{R}'\,\Omega_0$-Rule or $\mathrm{R}'\,\Omega_1$-Rule (if $l = 2$).

$\mathrm{R}'\,\Omega_1$-Rule. $t = c(\mathrm{R}'\,sab)$ for some context $c(\star)$ and there is some $\beta[\star] \in \mathcal{C}$ such that $\alpha = \beta[\omega] + 1$, $\vdash^{\beta[0]}_{\rho} s$ and for $u \in \mathrm{T}(\mathcal{V})$, $k < \omega$ with $\mathrm{lev}(u) = 0$ and $\models^{k} u$ also $\vdash^{\beta[k]}_{\rho} c(\mathrm{R}'\,uab)$, hence $\vdash^{1+\beta[k]}_{\rho} c(\mathrm{R}'\,uab)y$ by induction hypothesis. Let $\gamma[\star] := 1 + \beta[\star]$, then $\gamma[\star] \in \mathcal{C}$, $\gamma[\omega] + 1 = 1 + \beta[\omega] + 1 = 1 + \alpha$ and $\vdash^{\gamma[0]}_{\rho} s$ by Acc-Rule, hence $\vdash^{1+\alpha}_{\rho} ty$ by $\mathrm{R}'\,\Omega_1$-Rule. QED.

Collapsing Theorem 4.8 $\vdash^{\alpha}_{0} t \Rightarrow \models^{\mathcal{D}\alpha} \widehat{t}$

Proof. The proof is by induction on the definition of $\vdash^{\alpha}_{0} t$.

Acc-Rule. The assertion follows directly from the induction hypothesis and the fact that $\beta \leq_0 \alpha \Rightarrow \mathcal{D}\beta \leq \mathcal{D}\alpha$.

$\overline{\mathcal{V}}_{0,\mathrm{S},\mathrm{R}}$-Rule. $\mathrm{rr}(t) \in \overline{\mathcal{V}}_{0,\mathrm{S},\mathrm{R}}$, $\alpha = \beta + \#\,\mathrm{mat}(t)$ and $\forall s \in \mathrm{mat}(t) \vdash^{\beta}_{0} s$. We have $\mathrm{rr}(\widehat{t}) = \widehat{\mathrm{rr}(t)} \in \mathcal{V}_{0,\mathrm{S},\mathrm{R}}$ and $\mathrm{mat}(t) = \widehat{\mathrm{mat}(t)}$, thus $\models^{\mathcal{D}\beta} s$ for all $s \in \mathrm{mat}(\widehat{t})$ by induction hypothesis. As

$$\mathcal{D}\beta + \#\,\mathrm{mat}(\widehat{t}) = \mathcal{D}\beta + \#\,\mathrm{mat}(t) = \mathcal{D}(\beta + \#\,\mathrm{mat}(t)) = \mathcal{D}\alpha$$

we obtain $\models^{\mathcal{D}\alpha} t$ by $\mathcal{V}_{0,\mathrm{S},\mathrm{R}}$-Rule.

β-Rule. $\mathrm{rr}(t) = (\lambda\overline{x}r)s$, $\alpha = \beta + 1$ and $\mid\!\frac{\beta}{0}\ \mathrm{coat}_t(r[\overline{x} := s])$, $\mid\!\frac{\beta}{0}\ s$. We have $\mathrm{rr}(\widehat{t}) = \widehat{\mathrm{rr}(t)} = (\lambda x\widehat{r})\widehat{s}$, $(r[\overline{x} := s])\widehat{\ } = \widehat{r}[x := \widehat{s}]$ because if $\overline{x} = x'$ then x does not occur in r, and hence $\mathrm{coat}_t(r[\overline{x} := s])\widehat{\ } = \mathrm{coat}_{\widehat{t}}(\widehat{r}[x := \widehat{s}])$. By induction hypothesis we get $\models^{\mathcal{D}\beta} \mathrm{coat}_{\widehat{t}}(\widehat{r}[x := \widehat{s}])$ and $\models^{\mathcal{D}\beta} \widehat{s}$. As $\mathcal{D}\beta < \mathcal{D}\beta + 1 = \mathcal{D}\alpha$ we obtain $\models^{\mathcal{D}\alpha} \widehat{t}$ by β-Rule.

$\overline{\mathrm{R}}0$-Rule and $\overline{\mathrm{R}}\,\mathrm{S}$-Rule are similar to β-Rule.

Cut-Rule is not possible

$\mathrm{R}'\,\Omega_0$-Rule. $t = \mathrm{R}'\,u_1 \dots u_l$ with $l \leq 2$ and there is some $\gamma := \beta[0]$ with $\gamma + 2 \leq_0^* \alpha$ and $\mid\!\frac{\gamma}{0}\ u_i$ for $1 \leq i \leq l$. Then $\widehat{t} = \mathrm{R}\,\widehat{u_1} \dots \widehat{u_l}$, $\mathrm{rr}(\widehat{t}) = \mathrm{R}$ and $\mathrm{mat}(\widehat{t}) = \{\{\widehat{u_1}, \dots, \widehat{u_l}\}\}$. By induction hypothesis $\models^{\mathcal{D}\gamma} s$ for all $s \in \mathrm{mat}(\widehat{t})$ and hence $\models^{\mathcal{D}\alpha} \widehat{t}$ by $\mathcal{V}_{0,\mathrm{S},\mathrm{R}}$-Rule because $\mathcal{D}\gamma + l \leq \mathcal{D}(\gamma + 2) \leq \mathcal{D}\alpha$.

$\mathrm{R}'\,\Omega_1$-Rule. $t = c(\mathrm{R}'\,sab)$ and there is some $\beta[\star] \in \mathcal{C}$ such that $\alpha = \beta[\omega] + 1$, $\mid\!\frac{\beta[0]}{0}\ s$ and

$$\forall u \in \mathrm{T}(\mathcal{V}) \forall k < \omega \Big(\mathrm{lev}(u) = 0 \;\&\; \models^{k} u \Rightarrow \mid\!\frac{\beta[k]}{0}\ c(\mathrm{R}'\,uab) \Big) \tag{4}$$

With induction hypothesis we obtain $\models^{\mathcal{D}\beta[0]} \widehat{s}$. Now $\widehat{s} \in \mathrm{T}(\mathcal{V})$, $\mathrm{lev}(\widehat{s}) = 0$ and $\mathcal{D}\beta[0] < \omega$, thus

$$\mid\!\frac{\beta[\mathcal{D}\beta[0]]}{0}\ c(\mathrm{R}'\,\widehat{s}ab)$$

by (4). We have $(c(\mathrm{R}'\,\widehat{s}ab))\widehat{\ } = \widehat{c}(\mathrm{R}\,\widehat{s}\widehat{a}\widehat{b}) = \widehat{t}$, hence $\models^{\mathcal{D}\beta[\mathcal{D}\beta[0]]} \widehat{t}$ again by induction hypothesis. Now comes the highlight: $\mathcal{D}\beta[\mathcal{D}\beta[0]] < \mathcal{D}\beta[\omega] < \mathcal{D}\alpha$, hence $\models^{\mathcal{D}\alpha} \widehat{t}$.

QED.

Substitution Lemma 4.9 *$\mid\!\frac{\alpha}{\rho}\ r$, $\mid\!\frac{\beta}{\rho}\ s_j$, $\mathrm{lev}(s_j) \leq \rho$, $x_j \in \mathcal{V}'$, $s_j \in \mathrm{T}'(\mathcal{V})$ for $j < l$ then $\mid\!\frac{\beta+\alpha}{\rho}\ r[\vec{x} := \vec{s}]$.*

Proof. The proof is by induction on the definition of $\mid\!\frac{\alpha}{\rho}\ r$. Let u^* be $u[\vec{x} := \vec{s}]$.

Acc-Rule. The assertion follows directly from the induction hypothesis and the fact that $\gamma \leq_0 \alpha \Rightarrow \beta + \gamma \leq_0 \beta + \alpha$.

$\overline{\mathcal{V}}_{0,\mathrm{S},\mathrm{R}}$-Rule. $\mathrm{rr}(r) \in \overline{\mathcal{V}}_{0,\mathrm{S},\mathrm{R}}$ and there is some γ such that $\alpha = \gamma + \#\,\mathrm{mat}(r)$ and $\forall u \in \mathrm{mat}(r) \mid\!\frac{\gamma}{\rho}\ u$.

If $\mathrm{rr}(r) \notin \{\vec{x}\}$ then $\mathrm{rr}(r^*) = \mathrm{rr}(r)$ because $x_j \in \mathcal{V}'$ by assumption. We have $\mathrm{mat}(r^*) = \mathrm{mat}(r)^*$ and therefore $\mid\!\frac{\beta+\gamma}{\rho}\ u$ for all $u \in \mathrm{mat}(r^*)$ by induction hypothesis. Now $\#\,\mathrm{mat}(r^*) = \#\,\mathrm{mat}(r)$, hence $\beta + \gamma + \#\,\mathrm{mat}(r^*) = \beta + \alpha$, thus $\mid\!\frac{\beta+\alpha}{\rho}\ r^*$ by $\overline{\mathcal{V}}_{0,\mathrm{S},\mathrm{R}}$-Rule.

Now assume $\mathrm{rr}(r) = x_j$, then $r = x_j r_1 \dots r_n$, $n = \#\,\mathrm{mat}(r)$, and by induction hypothesis $\mid\!\frac{\beta+\gamma}{\rho}\ r_i^*$ for $1 \leq i \leq n$. From the assumptions we obtain $\mid\!\frac{\beta+\gamma}{\rho}\ s_j$ and $\mathrm{lev}(s_j) \leq \rho$. With Acc-Rule we receive $\mid\!\frac{\beta+\gamma+i-1}{\rho}\ r_i^*$ for $1 \leq i \leq n$. Therefore applying i cuts yields $\mid\!\frac{\beta+\gamma+i}{\rho}\ s_j r_1^* \dots r_i^*$, hence $\mid\!\frac{\beta+\alpha}{\rho}\ r^*$.

β-Rule, $\overline{\mathrm{R}}0$-Rule, $\overline{\mathrm{R}}\,\mathrm{S}$-Rule and Cut-Rule: The assertion follows directly from the induction hypothesis by applying the same inference.

R′ Ω_0-Rule. $r = \mathrm{R}'\, u_1 \dots u_l$ and $l \leq 2$, there are new variables, distinct in pairs, $u_{l+1}, \dots, u_3 \in \mathcal{V}'$ (w.l.o.g. they are also new for $\vec{x}$ and $\vec{s}$) and $\gamma[\star] \in \mathcal{C}$ such that $\alpha = \gamma[\omega] + 1$, $\gamma[0] + 2 \leq_0^* \alpha$, $\mid\!\frac{\gamma[0]}{\rho}\, u_i$ for $1 \leq i \leq l$ and

$$\forall u \in \mathrm{T}(\mathcal{V}) \forall k < \omega \Big(\mathrm{lev}(u) = 0 \;\&\; \models^{k} u \Rightarrow \mid\!\frac{\gamma[k]}{\rho}\; \mathrm{R}'\, u u_2 u_3 \Big)$$

We have $r^* = \mathrm{R}'\, u_1^* \dots u_l^*$ and $\beta + \gamma[\star] \in \mathcal{C}$ with $\beta + \gamma[\omega] + 1 = \beta + \alpha$, $\beta + \gamma[0] + 2 \leq_0^* \beta + \alpha$. By induction hypothesis $\mid\!\frac{\beta+\gamma[0]}{\rho}\, u_i^*$ for $1 \leq i \leq l$ and

$$\forall u \in \mathrm{T}(\mathcal{V}) \forall k < \omega \Big(\mathrm{lev}(u) = 0 \;\&\; \models^{k} u \Rightarrow \mid\!\frac{\beta+\gamma[k]}{\rho}\; \mathrm{R}'\, u u_2^* u_3^* \Big)$$

because for $u \in \mathrm{T}(\mathcal{V})$ $x_j \in \mathcal{V}'$ does not occur in u. Hence $\mid\!\frac{\beta+\alpha}{\rho}\, r^*$ by R′ Ω_0-Rule.
R′ Ω_1-Rule. $t = c(\mathrm{R}'\, sab)$ and there is some $\gamma[\star] \in \mathcal{C}$ such that $\alpha = \gamma[\omega] + 1$, $\mid\!\frac{\gamma[0]}{\rho}\, s$ and

$$\forall u \in \mathrm{T}(\mathcal{V}) \forall k < \omega \Big(\mathrm{lev}(u) = 0 \;\&\; \models^{k} u \Rightarrow \mid\!\frac{\gamma[k]}{\rho}\; c(\mathrm{R}'\, uab) \Big)$$

We have $r^* = c^*(\mathrm{R}'\, s^* a^* b^*)$ and $\beta + \gamma[\star] \in \mathcal{C}$ with $\beta + \gamma[\omega] + 1 = \beta + \alpha$. By induction hypothesis $\mid\!\frac{\beta+\gamma[0]}{\rho}\, s^*$ and

$$\forall u \in \mathrm{T}(\mathcal{V}) \forall k < \omega \Big(\mathrm{lev}(u) = 0 \;\&\; \models^{k} u \Rightarrow \mid\!\frac{\beta+\gamma[k]}{\rho}\; c^*(\mathrm{R}'\, u a^* b^*) \Big)$$

hence $\mid\!\frac{\beta+\alpha}{\rho}\, r^*$ by R′ Ω_1-Rule. QED.

Cut Elimination Lemma 4.10 $\mid\!\frac{\alpha}{\rho+1}\, t \Rightarrow \mid\!\frac{2^\alpha}{\rho}\, t$

We cannot prove this Lemma in this formulation by induction on the definition of $\mid\!\frac{\alpha}{\rho+1}\, t$, because cuts are replaced by appending a variable and afterwards applying the Substitution Lemma which leads to the sum of the derivation lengths plus 1. Thus we would need $2^\beta + 2^\beta + 1 \leq_0^* 2^{\beta+1}$ which is only true if we interpret the formal term 2^α by some ordinal function $3^{\mathcal{O}(\beta)+1}$ which we do not want.

We will need the following estimations

$$n < \omega \Rightarrow n + 1 \leq_0^* 2^n \tag{5}$$

$$\beta \neq 0, 0 < n < \omega \Rightarrow n + 1 + 2^\beta \leq_0^* 2^{\beta+n} \tag{6}$$

which can be proved by induction on n: $0+1 \leq_0^* 2^0$, and by induction hypothesis $k + 1 \leq_0^* 2^k$, hence $(k+1) + 1 \leq_0^* 2^k + 1 \leq_0 2^k + 2^k \leq_0 2^{k+1}$. Using (5) we obtain $1 + 1 + 2^\beta \leq_0 2^1 + 2^\beta \leq_0 2^\beta + 2^\beta \leq_0 2^{\beta+1}$. By induction hypothesis $k + 1 + 2^\beta \leq_0^* 2^{\beta+k}$, hence $(k+1) + 1 + 2^\beta \leq_0^* 1 + 2^{\beta+k} \leq_0^* 2^{\beta+k+1}$.
Proof of the Cut Elimination Lemma. We show by induction on the definition of $\mid\!\frac{\alpha}{\rho+1}\, t$

$$\mid\!\frac{\alpha}{\rho+1}\, t \Rightarrow \exists \beta (1 + \beta \leq_0^* 2^\alpha \;\&\; \mid\!\frac{\beta}{\rho}\, t).$$

Then the main assertion simply follows by a Structural Rule.

Acc-Rule. The assertion follows directly from the induction hypothesis and the fact that $\gamma \leq_0 \alpha \Rightarrow 2^\gamma \leq_0 2^\alpha$ and therefore $1+\beta \leq_0^* 2^\gamma \Rightarrow 1+\beta \leq_0^* 2^\alpha$.

$\overline{\mathcal{V}}_{0,\mathrm{S},\mathrm{R}}$-Rule. $\mathrm{rr}(t) \in \overline{\mathcal{V}}_{0,\mathrm{S},\mathrm{R}}$, $\alpha = \beta + \#\,\mathrm{mat}(t)$ and $\forall s \in \mathrm{mat}(t) \mathrel{|\!\frac{\beta}{\rho+1}} s$. Let $n := \#\,\mathrm{mat}(t)$.

If $n = 0$ then $\mathrel{|\!\frac{0}{\rho}} t$ and $1+0 \leq_0 2^\alpha$. If $\beta = 0$ then $\forall s \in \mathrm{mat}(t) \mathrel{|\!\frac{0}{\rho}} s$, thus $\mathrel{|\!\frac{n}{\rho}} t$. Now $n+1 \leq_0^* 2^n$ by (5).

Otherwise $\beta \neq 0$ and $n = n'+1$. By induction hypothesis we obtain $\forall s \in \mathrm{mat}(t) \mathrel{|\!\frac{2^\beta}{\rho}} s$, thus $\mathrel{|\!\frac{2^\beta+n}{\rho}} t$ and $1+2^\beta+n \leq_0^* 1+n+2^\beta \leq_0^* 2^{\beta+n}$ by (6).

β-Rule. $\mathrm{rr}(t) = (\lambda x r)s$, $\alpha = \beta+1$ and $\mathrel{|\!\frac{\beta}{\rho+1}} \mathrm{coat}_t(r[x:=s])$, $\mathrel{|\!\frac{\beta}{\rho+1}} s$.

If $\beta = 0$ then $\mathrel{|\!\frac{0}{\rho}} \mathrm{coat}_t(r[x:=s])$, $\mathrel{|\!\frac{0}{\rho}} s$, hence $\mathrel{|\!\frac{1}{\rho}} t$ by β-Rule and we have $1+1 \leq_0^* 2^1$ by (5).

Now assume $\beta \neq 0$, then by induction hypothesis $\mathrel{|\!\frac{2^\beta}{\rho}} \mathrm{coat}_t(r[x:=s])$, $\mathrel{|\!\frac{2^\beta}{\rho}} s$, hence $\mathrel{|\!\frac{2^\beta+1}{\rho}} t$ and $1+2^\beta+1 \leq_0 1+1+2^\beta \leq_0^* 2^{\beta+1}$ by (6).

$\overline{\mathrm{R}}0$-Rule and $\overline{\mathrm{R}}\,\mathrm{S}$-Rule are similar to β-Rule.

Cut-Rule. $r = (st)$, $\mathrm{lev}(s) \leq \rho+1$, $t \in \mathrm{T}'(\mathcal{V})$, $\alpha = \beta+1$ and $\mathrel{|\!\frac{\beta}{\rho+1}} s$ and $\mathrel{|\!\frac{\beta}{\rho+1}} t$. By induction hypothesis there are γ_1, γ_2 with $1+\gamma_i \leq_0^* 2^\beta$ and $\mathrel{|\!\frac{\gamma_1}{\rho}} s$ and $\mathrel{|\!\frac{\gamma_2}{\rho}} t$. The Appending Lemma shows $\mathrel{|\!\frac{2^\beta}{\rho}} sy$ for some $y \in \mathcal{V}'$, thus $\mathrel{|\!\frac{\gamma_2+2^\beta}{\rho}} r$ by the Substitution Lemma as $\mathrm{lev}(t) \leq \rho$. We compute $1+\gamma_2+2^\beta \leq_0^* 2^\beta+2^\beta \leq_0 2^{\beta+1}$.

$\mathrm{R}'\,\Omega_0$-Rule, $\mathrm{R}'\,\Omega_1$-Rule: By induction hypothesis we obtain $\mathrel{|\!\frac{2^{\beta[\omega]}+1}{\rho}} t$ for some $\beta[\star] \in \mathcal{C}$ with $\beta[\omega]+1 = \alpha$, because we also have $2^{\beta[\star]} \in \mathcal{C}$. Now $1+2^{\beta[\omega]}+1 \leq_0 1+1+2^{\beta[\omega]} \leq_0^* 2^{\beta[\omega]+1}$ by (6). QED.

Lemma 4.11 *Let* $\mathrm{R}'\,0ab \in \overline{\mathrm{T}}(\mathcal{V})$ *with variables* $a, b \in \mathcal{V}'$ *and let* $\rho = \mathrm{lev}(a)$, *then*

$$\models^{\alpha} t \ \text{ and } \ \mathrm{lev}(t) = 0 \Rightarrow \mathrel{|\!\frac{2+2\cdot\alpha}{\rho}} \mathrm{R}'\,tab.$$

Proof. The proof is by induction on the definition of $\models^{\alpha} t$.

$\mathcal{V}_{0,\mathrm{S},\mathrm{R}}$-Rule. $\mathrm{rr}(t) \in \mathcal{V}_{0,\mathrm{S},\mathrm{R}}$, $M := \mathrm{mat}(t)$, $n := \#M$ and there is some β such that $\beta+n \leq \alpha$ and $\forall s \in M \models^{\beta} s$.

If $\mathrm{rr}(t) = 0$ then $t = 0$. We have $\mathrel{|\!\frac{0}{\rho}} a$, $\mathrel{|\!\frac{0}{\rho}} b$, hence $\mathrel{|\!\frac{1}{\rho}} \mathrm{R}'\,tab$ by $\overline{\mathrm{R}}0$-Rule.

If $\mathrm{rr}(t) = \mathrm{S}$ then $t = \mathrm{S}\,t'$, hence $n = 1$ and $\models^{\beta} t'$. Let $\gamma = 2\cdot\beta+1$, then $\mathrel{|\!\frac{\gamma}{\rho}} t'$ by the subsystem property. Now $\mathrel{|\!\frac{\gamma}{\rho}} a$ by the $\mathcal{V}_{0,\mathrm{S},\mathrm{R}}$-Rule, hence $\mathrel{|\!\frac{\gamma+1}{\rho}} at'$ by the Cut-Rule as $\mathrm{lev}(a) = \rho$. The induction hypothesis yields $\mathrel{|\!\frac{\gamma+1}{\rho}} \mathrm{R}'\,t'ab$, thus again applying the Cut-Rule produces $\mathrel{|\!\frac{\gamma+2}{\rho}} at'(\mathrm{R}'\,t'ab)$ as $\mathrm{lev}(at') \leq \mathrm{lev}(a) = \rho$. Thus $\mathrel{|\!\frac{\gamma+3}{\rho}} \mathrm{R}'\,tab$ using the $\overline{\mathrm{R}}\,\mathrm{S}$-Rule, and $\gamma+3 = 2+2\cdot(\beta+1) \leq 2+2\cdot\alpha$.

$\mathrm{rr}(t) = \mathrm{R}$ is not possible because $\mathrm{lev}(t) = 0$.

It remains $\mathrm{rr}(t) \in \mathcal{V}$, thus $\mathrm{rr}(\mathrm{R}'\,tab) = \mathrm{rr}(t) \in \mathcal{V}$ and $\mathrm{mat}(\mathrm{R}'\,tab) = M \cup \{\{a,b\}\}$. By the subsystem property we have $\forall s \in M \mathrel{|\!\frac{\beta}{\rho}} s$, as well as $\mathrel{|\!\frac{0}{\rho}} a$, $\mathrel{|\!\frac{0}{\rho}} b$, thus $\mathrel{|\!\frac{\beta+n+2}{\rho}} \mathrm{R}'\,tab$ by $\overline{\mathcal{V}}_{0,\mathrm{S},\mathrm{R}}$-Rule. Now $\beta+n+2 \leq 2+\alpha$.

β-Rule. $\mathrm{rr}(t) = (\lambda xr)s$ and there is some $\beta < \alpha$ such that $\overset{\beta}{\models} \mathrm{coat}_t(r[x := s])$ and $\overset{\beta}{\models} s$. We have $\mathrm{redex}(t) = (\lambda xr)s$ because $\mathrm{lev}(t) = 0$. By induction hypothesis $\vdash^{\gamma}_{\rho} \mathrm{R}' \, \mathrm{coat}_t(r[x := s])ab$ for $\gamma = 2 + 2 \cdot \beta$. The subsystem property shows $\overset{\beta}{\models} s \Rightarrow \vdash^{\beta}_{0} s$, hence $\vdash^{\gamma}_{\rho} s$. Thus $\vdash^{\gamma+1}_{\rho} \mathrm{R}' \, tab$ by β-Rule.
R 0-Rule and R S-Rule are similar to β-Rule. QED.

The *length* $\mathrm{l}(r)$ of r is defined by $\mathrm{l}(x) = 1$, $\mathrm{l}(\lambda xr) = \mathrm{l}(r)+1$, $\mathrm{l}(rs) = \mathrm{l}(r)+\mathrm{l}(s)$, and the *height* $\mathrm{h}(r)$ by $\mathrm{h}(x) = 0$, $\mathrm{h}(\lambda xr) = \mathrm{h}(r)+1$, $\mathrm{h}(rs) = \max(\mathrm{h}(r), \mathrm{h}(s))+1$. By induction on r we immediately see $\mathrm{l}(r) \leq 2^{\mathrm{h}(r)}$.

Embedding Lemma 4.12 *$t \in \mathrm{T}(\mathcal{V})$ and $\mathrm{g}(t) \leq \rho + 1 \Rightarrow \vdash^{2^{\omega+1}\cdot \mathrm{l}(t)}_{\rho} \bar{t}$.*

Proof. . Let $e(k) := 4 \cdot k - 1 + 2^{\omega} \cdot k$ for $k > 0$. Then $e(k) \leq^*_0 2^2 \cdot k + 2^{\omega} \cdot k \leq^*_0 (2^{\omega} + 2^{\omega}) \cdot k \leq^*_0 2^{\omega+1} \cdot k$. We prove

$$\mathrm{g}(t) \leq \rho + 1 \Rightarrow \vdash^{e(\mathrm{l}(t))}_{\rho} \bar{t}$$

by induction on the definition of $t \in \mathrm{T}(\mathcal{V})$, then the assertion follows by a Structural Rule.
$t \in \mathcal{V}_{0,\mathrm{S}}$. We have $\vdash^{0}_{0} \bar{t}$ by $\overline{\mathcal{V}}_{0,\mathrm{S},\mathrm{R}}$-Rule.
$t = \mathrm{R}$. Let $a, b \in \mathcal{V}'$ such that $\mathrm{R}' \, 0ab \in \mathrm{T}'(\mathcal{V})$, then the previous Lemma shows

$$\forall u \in \mathrm{T}(\mathcal{V}) \forall k < \omega \Big(\mathrm{lev}(u) = 0 \;\&\; \overset{k}{\models} u \Rightarrow \vdash^{2+2\cdot k}_{\rho} \mathrm{R}' \, uab \Big)$$

because $\mathrm{lev}(a) < \mathrm{lev}(\mathrm{R}') \leq \rho+1$. Setting $\beta[\star] := 2+2^{\star} \in \mathcal{C}$ we obtain $2+2\cdot k \leq^*_0 \beta[k]$ by induction on k, where $2 \leq_0 \beta[0]$ and $4 \leq^*_0 \beta[1]$ are clear, and for $k > 0$ with induction hypothesis $2+2\cdot(k+1) \leq^*_0 \beta[k]+1+1 \leq^*_0 2+2^k+2^k \leq_0 2+2^{k+1} = \beta[k+1]$. Furthermore $\beta[0]+2 = 2+2^0+1+1 \leq^*_0 2+2^1+1 \leq_0 2+2^{\omega}+1 = \beta[\omega]+1$ and $\beta[\omega] + 1 = 2 + 2^{\omega} + 1 \leq_0 3 + 2^{\omega} = e(1)$, hence $\vdash^{e(1)}_{\rho} \mathrm{R}'$ by $\mathrm{R}' \, \Omega_0$-Rule and a Structural Rule.
$t = \lambda xr$. Then $\mathrm{g}(r) \leq \rho + 1$, hence $\vdash^{e(\mathrm{l}(r))}_{\rho} \bar{r}$ by induction hypothesis. Hence $\vdash^{e(\mathrm{l}(r)+1)}_{\rho} \bar{t}$ by β-Rule.
$t = (rs)$. Then $\mathrm{g}(r), \mathrm{g}(s) \leq \rho + 1$, hence $\vdash^{e(\mathrm{l}(r))}_{\rho} \bar{r}$ and $\vdash^{e(\mathrm{l}(s))}_{\rho} \bar{s}$ by induction hypothesis. The Appending Lemma shows $\vdash^{e(\mathrm{l}(r))+1}_{\rho} \bar{r}z$ for some suitable $z \in \mathcal{V}'$, hence $\vdash^{e(\mathrm{l}(s))+e(\mathrm{l}(r))+1}_{\rho} \bar{r}\,\bar{s}$ using the Substitution Lemma, because $\mathrm{lev}(\bar{s}) < \mathrm{lev}(\bar{r}) \leq \rho+1$ and $\bar{s} \in \mathrm{T}'(\mathcal{V})$. Now $e(m)+e(n)+1 \leq^*_0 4\cdot(m+n)-1+2^{\omega}\cdot(m+n) = e(m+n)$, hence $\vdash^{e(\mathrm{l}(t))}_{\rho} \bar{t}$. QED.

Now we put everything together. Let $t \in \mathrm{T}(\mathcal{V})$ with $\mathrm{g}(t) = \rho + 1$. The Embedding Lemma and the Cut Elimination Lemma show

$$\vdash^{2_{\rho}(2^{\omega+1}\cdot \mathrm{l}(t))}_{0} \bar{t}$$

where $2_n(\alpha)$ is the obvious term defined by iteration of 2^{α}, i.e. $2_0(\alpha) = \alpha$ and $2_{n+1}(\alpha) = 2^{2_n(\alpha)}$. Now the Collapsing Theorem leads to

$$\overset{\mathcal{D}2_{\rho}(2^{\omega+1}\cdot \mathrm{l}(t))}{\models} t$$

because $\widehat{\widehat{t}} = t$. Hence we obtain with the Estimate Theorem

$$\begin{aligned} \mathrm{d}(t) &\le 2^{\mathcal{D}2_\rho(2^{\omega+1}\cdot \mathrm{l}(t))} = 2^{\Psi(\mathcal{O}(2_\rho(2^{\omega+1}\cdot \mathrm{l}(t)))} \\ &\le \begin{cases} 2^{\Psi(\omega\cdot 4\cdot \mathrm{l}(t))} & : \rho = 0 \\ 2^{\Psi(w_\rho(4\cdot \mathrm{l}(t)+1))} & : \rho > 0 \end{cases} \\ &\le \begin{cases} 2^{\Psi(\omega\cdot 4\cdot 2^{\mathrm{h}(t)})} & : \rho = 0 \\ 2^{\Psi(w_\rho(4\cdot 2^{\mathrm{h}(t)}+1))} & : \rho > 0 \end{cases} \end{aligned}$$

It follows from [S97] and [BCW94] that these bounds are optimal.

Remark 4.13 GÖDEL*'s T in the formulation with combinators K and S can also be analyzed using the same machinary from this paper obtaining the same results. To this end we have to replace the β-Rules by rules for K and S. They are treated similiar to the recursor, of course without Ω-rules, but also with copies K' and S' for handling substitution, i.e. cut-elimination.*

References

[Be98] Beckmann, A.: *Exact bounds for lengths of reductions in typed λ-calculus.* submitted to JSL, Münster (1998)

[Bu80] Buchholz, W.: Three contributions to the conference on recent advances in proof theory Oxford 1980, mimeographed.

[BCW94] Buchholz, W., E.A. Cichon, and Weiermann A.: *A uniform approach to fundamental sequences and hierarchies*, Mathematical Logic Quarterly 40 (1994), 273-286.

[S97] Schwichtenberg, H.: *Classifying recursive functions.* Draft for the Handbook of Recursion Theory (ed. E. Griffor) (1997) `http://www.mathematik.uni-muenchen.de/~schwicht/`

[S91] Schwichtenberg, H.: *An upper bound for reduction sequences in the typed λ-calculus.* Arch. Math. Logic 30, 405–408 (1991)

[S82] Schwichtenberg, H.: *Complexity of normalization in the pure typed lambda - calculus.* In: The L.E.J.Brouwer Centenary Symposium, A.S. Troelstra and D. van Dalen (editors), North Holland, 453–457 (1982)

[W98] Weiermann, A.: *How is it that infinitary methods can be applied to finitary mathematics? Gödel's T: a case study* Journal of Symbolic Logic 63 (1998), 1348-1370.

[WW98] Wilken, G.,Weiermann, A.:*Sharp upper bounds for the depths of reduction trees of typed λ – calculus with recursors* (submitted).