Reasonable Bounds for Lengths of Reduction Chains in Acyclic HORS

Dominik Wehr, ILLC, University of Amsterdam
Supervised*by Bahareh Afshari, ILLC, University of Amsterdam

July 10, 2021

In his paper of 2001 [1], Arnold Beckmann derived bounds on reduction lengths for the simply typed lambda calculus. We adapt his approach to derive similar bounds for acyclic higher-order reduction schemes with a non-deterministic choice operator. Furthermore, we derive language size bounds for AHORS based said reduction chain bounds.

1 Outline

While the details of these results are quite technical, the general approach to deriving them is strikingly elegant. We thus start by outlining the results and proof strategies to obtain them.

1.1 Reduction Chain Length Bounds

We derive uniform bounds on the lengths of reduction chains for certain classes of terms of acyclic HORS. These classes are characterized by the degree $\deg(t)$ (the highest order among the recursive subterms of t), the recursive height $h(t)$ or the recursive size $s(t)$.

$$\begin{aligned} ds_n(N) &\coloneqq \max\{d(t) \mid \deg(t) \leq n \text{ and } s(t) \leq N\} \\ dh_n(N) &\coloneqq \max\{d(t) \mid \deg(t) \leq n \text{ and } h(t) \leq N\} \end{aligned}$$

The overall results we arrive at are that for any fixed n we have

$$ds_n(N) = 2_n(\Theta(N)) \qquad dh_n(N) = 2_{n+1}(\Theta(N))$$

*This report was created in the course of a Master of Logic individual project where $2_0(N) = N$ and $2_{n+1}(N) = 2^{2_n(N)}$. More concretely, this yields

$$ds_n(N) \leq 2_n(N) \qquad dh_n(N) \leq 2_{n+1}(N)$$

the proof of these results is split into two steps.

Deriving the upper bound We use expanded head reduction trees, formalized as a derivation system $\Downarrow^\alpha_\rho s$. $\Downarrow^\alpha_\rho s$ states that there exists such a tree of height α which "accounts for" any all future δ-transition steps of non-terminals of an order greater than ρ which can occur within any term $s \rightsquigarrow^* s'$. To derive the upper bound, we then combine three ideas:

1. We have $d(s) \leq |s| \leq 2^\alpha$ where s denotes the size of the tree $\Downarrow^\alpha_0$, ie. the tree accounting for *all* future δ-transitions in s.
2. For any $\Downarrow^\alpha_\rho s$ we may reduce ρ by one at the cost of increasing the height α exponentially. This means $\Downarrow^\alpha_{\rho+1} s$ implies $\Downarrow^{2^\alpha}_\rho s$.
3. For a well-typed acyclic HORS term t we have $\Downarrow^{s(t)}_{\deg(t)-1} t$ and $\Downarrow^{h(t)}_{\deg(t)} t$.

Which together yield

$$d(t) \leq 2_{\deg(t)+1}(h(t)) \quad \text{and} \quad d(t) \leq 2_{\deg(t)}(s(t))$$

as well as

$$dh_n(N) = 2_{n+1}(O(N)) \quad \text{and} \quad ds_n(N) = 2_n(O(N))$$

Deriving the lower bound We build AHORS terms $D^s_n(N)$ and $D^h_n(N)$ with $s(D^s_n(N)) = O(N) = h(h^N_n)$, $\deg(D^s_n(N)) = \deg(D^h_n(N)) = n$ and $d(D^s_n(N)) \geq 2_n(N)$ as well as $d(D^h_n(N)) \geq 2_{n+1}$. These yield

$$dh_n(N) = 2_{n+1}(\Omega(N)) \quad \text{and} \quad ds_n(N) = 2_n(\Omega(N))$$

1.2 Language Size Bounds

We derive similarly uniform bounds on the size of the language generated by an AHORS term. For this, we first define size- and height-based measures.

$$l(t) := |\{s \mid t \rightsquigarrow^* s \text{ and } s \text{ normal}\}|$$
$$ls_n(N) := \max\{l(t) \mid \deg(t) \leq n \text{ and } s(t) \leq N\}$$
$$lh_n(N) := \max\{l(t) \mid \deg(t) \leq n \text{ and } h(t) \leq N\}$$

The results we derive are as follows

$$ls_n(N) \leq 2_{n+1}(\Theta(N)) \qquad lh_n(N) \leq 2_{n+2}(\Theta(N))$$

The proof of this result again proceeds via two separate steps.

Deriving the upper bound This result is obtained by a more fine-grained analysis of the nature of possible transitions chains of AHORS. We distinguish two different kinds of reduction steps: The unfolding of non-terminals $\rightsquigarrow_\delta$ and the reduction of the non-deterministic choice operator $\rightsquigarrow_\gamma$. The upper bound is then obtained by combining a few different lemmas about them:

1. The $\rightsquigarrow_\delta$ transitions confluent, meaning AHORS terms have a unique δ-normal form.
2. For a given AHORS term s, any of its normal forms t can be obtained via a reduction chain $s \rightsquigarrow_\delta^* s' \rightsquigarrow_\gamma^* t$ where s' is the δ-normal form of s.
3. We define $l_\gamma(t) := |\{t' \mid t \rightsquigarrow_\gamma^* t' \text{ and } t' \text{ is } \gamma\text{-normal}\}|$ and show that $l_\gamma(t) \leq 2^{d(t)}$.

Taken together, this yields

$$l(s) = l(s') = l_\gamma(s') \leq 2^{d(s')} \leq 2^{d(s)}$$

which can be combined with the reduction chain bounds already obtained to conclude

$$ls_n(N) \leq 2_{n+2}(N) = 2_{n+2}(O(N)) \qquad lh_n(N) \leq 2_{n+1}(N) = 2_{n+1}(O(N))$$

Deriving the lower bound Our approach directly parallels that for the reduction chain bounds. We build AHORS terms $L_n^s(N)$ and $L_n^h(N)$ with $s(L_n^s(N)) = O(N) = h(h_n^N)$, $\deg(L_n^s(N)) = \deg(L_n^h(N)) = n$ and $l(L_n^s(N)) \geq 2_{n+1}(N)$ as well as $l(L_n^h(N)) \geq 2_{n+2}(N)$. These yield

$$lh_n(N) = 2_{n+2}(\Omega(N)) \quad \text{and} \quad ls_n(N) = 2_{n+1}(\Omega(N))$$

2 Acyclic HORS with Nondeterminism

The reduction system we examine are acyclic higher-order reduction systems. The syntax of HORS expressions and their types is given below.

$$\begin{aligned} s, t \in \mathrm{Tm} &::= x \mid a \mid s\,t \mid s + t \\ \sigma, \tau \in \mathrm{Ty} &::= 0 \mid \sigma \to \tau \end{aligned}$$

Similarly to many programming languages, HORS allow for letters a to be defined as shorthands for a more complex HORS expression. We write $a : \tau := x_1, \ldots, x_n \mapsto t^a$ to denote that a is such a shorthand for the expression t^a which may refer to the parameters $x_1, \ldots, x_n$. Clearly, some choices of shorthands, such as $a : 0 := a$, allow for expressions whose reduction sequences diverge, making our question of size- and height-based reduction sequence bounds trivial. We thus restrict our attention to acyclic HORS, meaning those whose "call-graph" (ie. the graph tracking which shorthands refer to which other shorthands) is acyclic.

We begin by defining what it means for a HORS expression to be both acyclic and well-typed with the inductively defined judgment $\mathcal{N}, \mathcal{T}, \Gamma \vdash s : \tau$. Here, $\mathcal{N}$ is a context of typed non-terminals of the shape $a : \tau \coloneqq \vec{x} \mapsto t^a$ as explained above. $\mathcal{T}$ and Γ list typed terminals and variables with the syntax $a : \tau$ and $x : \tau$, respectively.

$$\textsc{Var}\ \frac{x : \tau \in \Gamma}{\mathcal{N}; \mathcal{T}; \Gamma \vdash x : \tau} \qquad \textsc{Term}\ \frac{a : \tau \in \mathcal{T}}{\mathcal{N}; \mathcal{T}; \Gamma \vdash a : \tau}$$

$$\textsc{Def}\ \frac{\mathcal{N}; \mathcal{T}; \mathrm{args}(\vec{x}, \tau) \vdash t : 0 \qquad n = \mathrm{aty}(\tau)}{\mathcal{N}, (a : \tau \coloneqq x_1, \ldots, x_n \mapsto t), \mathcal{N}'; \mathcal{T}; \Gamma \vdash a : \tau}$$

$$\textsc{App}\ \frac{\mathcal{N}; \mathcal{T}; \Gamma \vdash s : \sigma \to \tau \qquad \mathcal{N}; \mathcal{T}; \Gamma \vdash t : \sigma}{\mathcal{N}; \mathcal{T}; \Gamma \vdash s\, t : \tau}$$

$$\textsc{Sum}\ \frac{\mathcal{N}; \mathcal{T}; \Gamma \vdash s : 0 \qquad \mathcal{N}; \mathcal{T}; \Gamma \vdash t : 0}{\mathcal{N}; \mathcal{T}; \Gamma \vdash s + t : 0}$$

$$\mathrm{args}(\varepsilon, 0) \coloneqq \varepsilon \qquad \mathrm{args}(x_0\vec{x}, \sigma \to \tau) \coloneqq (x_0 : \sigma)\mathrm{args}(\vec{x}, \tau)$$

$$\mathrm{aty}(\sigma \to \tau) \coloneqq \mathrm{aty}(\tau) + 1 \qquad \mathrm{aty}(0) \coloneqq 0$$

There are multiple things worth noting about these definitions: First of all, observe that $\mathcal{N}; \mathcal{T}; \Gamma \vdash s : \tau$ also asserts the acyclicty of s: Whenever the Def-rule is applied, all shorthand definitions above and including the one currently being unfolded are removed from $\mathcal{N}$. This ensures that $\mathcal{N}$ embodies a strict linear order $<$ on the non-terminals such that $b < a$ for any non-terminal b occurring (recursively) in t^a. Such a linear ordering exists if and only if the call-graph is acyclic.

Secondly, note that whenever the Def-rule is applied, the variable context Γ is replaced with the argument variables of the non-terminal which is being unfolded. This is because no expression associated with a non-terminal may refer to variables outside of its arguments.

Lastly, observe that $\mathcal{N}; \mathcal{T}; \Gamma \vdash s : \tau$ only ensures the well-typedness of all non-terminals occurring recursively in s. It is possible that there is $(a : \tau \coloneqq \vec{x} \mapsto t^a)$ with $\mathcal{N}; \mathcal{T}; \mathrm{args}(\vec{x}, \tau) \nvdash t^a : 0$ which simply does not occur recursively in s and is thus never "checked" by $\mathcal{N}; \mathcal{T}; \Gamma \vdash s : \tau$. For our purposes, this does not spell any trouble as we are only interested in those non-terminals which may occur along the reduction sequence of s, which are precisely those occurring recursively in s.

Formally, we define the **size $s(t)$ and height $h(t)$** on the derivation of $\mathcal{N}; \mathcal{T}; \Gamma \vdash t : \tau$. This is well-defined as long as each the terminals and non-terminals are distinct and only one unfolding for each non-terminal is contained in $\mathcal{N}$ as under these circumstances the derivation of $\mathcal{N}; \mathcal{T}; \Gamma \vdash t : \tau$ (if it exists) is unique. We thus define

$$
\begin{aligned}
s(x) &= 0 & h(x) &= 0 & & \\
s(a) &= 0 & h(a) &= 0 & & \text{if } a \in \mathcal{T} \\
s(a) &= 1 + s(t^a) & h(a) &= 1 + h(a) & & \text{if } a \coloneqq \vec{x} \mapsto t^a \in \mathcal{N} \\
s(t\,t') &= 1 + s(t) + s(t') & h(t\,t') &= = 1 + \max\{h(t), h(t')\} & & \\
s(t + t') &= 1 + s(t) + s(t') & h(t + t') &= 1 + \max\{h(t), h(t')\} & &
\end{aligned}
$$

Furthermore, we define the **order $\mathrm{ord}(\tau)$** recursively as

$$\mathrm{ord}(\sigma \to \tau) \coloneqq \max\{\mathrm{ord}(\sigma) + 1, \mathrm{ord}(\tau)\} \qquad \mathrm{ord}(0) = 0$$

If $t : \tau$ is clear from the context, we also sometimes write $\mathrm{ord}(t)$ for $\mathrm{ord}(\tau)$. We write $\deg(s)$ for the **degree of s** which is the maximal order of types occurring in $\mathcal{N}; \mathcal{T}; \Gamma \vdash s : \tau$.

We close this section by defining the **reduction steps $\mathcal{N} \vdash s \rightsquigarrow s'$** of the HORS expressions. Observe that the context $\mathcal{N}$ of non-terminals is needed to determine what term t^a any given non-terminal a should be unfolded into. We usually omit $\mathcal{N}$, simply writing $s \rightsquigarrow s'$, if $\mathcal{N}$ is clear from the context. Furthermore, note that a non-terminal may only be unfolded when all of its arguments are present. Lastly, observe the expression $s_0 + s_1$ denotes the non-deterministic choice between the terms s_0 and s_1.

$$\frac{i \in \{0,1\}}{\mathcal{N} \vdash s_0 + s_1 \rightsquigarrow s_i} \qquad \frac{(a : \tau \coloneqq x_1, \ldots, x_n \mapsto t) \in \mathcal{N} \qquad n = \mathrm{aty}(\tau)}{\mathcal{N} \vdash a\, s_1 \ldots s_n \rightsquigarrow t[\vec{s}/\vec{x}]}$$

$$\gamma\,\frac{i \in \{0,1\} \quad \mathcal{N} \vdash s_i \rightsquigarrow s_i' \quad s_{1-i} = s_{1-i}'}{\mathcal{N} \vdash s_0 + s_1 \rightsquigarrow s_0' + s_1'} \qquad \delta\,\frac{i \in \{0,1\} \quad \mathcal{N} \vdash s_i \rightsquigarrow s_i' \quad s_{1-i} = s_{1-i}'}{\mathcal{N} \vdash s_0\, s_1 \rightsquigarrow s_0'\, s_1'}$$

In Sections 5 and 6 we distinguish two kinds of reduction steps: If the derivation of $s \rightsquigarrow t$ ends in an application of the γ-rule, we call it a γ-reduction and denote it by $s \rightsquigarrow_\gamma t$. Similarly, an $s \rightsquigarrow t$ whose derivation ends in an application of the δ-rule is called a δ-reduction.

3 Expanded Head Reduction Trees

The key tool for establishing the upper bounds on reduction sequences are **expanded head reduction trees**. These are derivation trees of judgments of the form $\mathcal{N}; \mathcal{T}; \Gamma \Downarrow^\alpha_\rho t : \tau$. Intuitively, such a judgment expresses that the term t of type τ is strongly normalizing.

The parameter α should be viewed as an upper bound of the depth of the proof tree. The parameter ρ restricts which instances of the Cut-rule may be used. Crucially, $\rho = 0$ indicates a judgment in the Cut-free fragment of the derivation system.

$$\textsc{Var}\frac{x : \tau_1 \to \ldots \to \tau_n \to \tau \in \Gamma \qquad \mathcal{N};\mathcal{T};\Gamma \Downarrow^{\alpha}_{\rho} t_1 : \tau_1 \quad \ldots \quad \mathcal{N};\mathcal{T};\Gamma \Downarrow^{\alpha}_{\rho} t_n : \tau_n}{\mathcal{N};\mathcal{T};\Gamma \Downarrow^{\alpha+n}_{\rho} x\, t_1 \ldots t_n : \tau}$$

$$\textsc{Term}\frac{a : \tau_1 \to \ldots \to \tau_n \to \tau \in \mathcal{T} \qquad \mathcal{N};\mathcal{T};\Gamma \Downarrow^{\alpha}_{\rho} t_1 : \tau_1 \quad \ldots \quad \mathcal{N};\mathcal{T};\Gamma \Downarrow^{\alpha}_{\rho} t_n : \tau_n}{\mathcal{N};\mathcal{T};\Gamma \Downarrow^{\alpha+n}_{\rho} a\, t_1 \ldots t_n : \tau}$$

$$\gamma\frac{\mathcal{N};\mathcal{T};\Gamma \vdash s_0 : 0 \quad \mathcal{N};\mathcal{T};\Gamma \Downarrow^{\alpha}_{\rho} s_0 : 0 \qquad \mathcal{N};\mathcal{T};\Gamma \vdash s_1 : 0 \quad \mathcal{N};\mathcal{T};\Gamma \Downarrow^{\alpha}_{\rho} s_1 : 0}{\mathcal{N};\mathcal{T};\Gamma \Downarrow^{\alpha+1}_{\rho} s_0 + s_1 : 0}$$

$$\delta\frac{\begin{array}{c}\mathcal{N};\mathcal{T};\Gamma, \mathrm{args}((x_{n+1}, \ldots, x_m), \tau) \Downarrow^{\alpha}_{\rho} t^a[\vec{t}/x_1, \ldots, x_n] : 0 \\ \mathcal{N};\mathcal{T};\Gamma \Downarrow^{\alpha}_{\rho} t_1 : \tau_1 \quad \ldots \quad \mathcal{N};\mathcal{T};\Gamma \Downarrow^{\alpha}_{\rho} t_n : \tau_n \\ (a : \tau_1 \to \ldots \to \tau_n \to \tau \coloneqq x_1, \ldots, x_n, \ldots, x_m \mapsto t^a) \in \mathcal{N} \\ \mathcal{N};\mathcal{T};\mathrm{args}(\tau_1 \to \ldots \to \tau, \vec{x}) \vdash t^a : 0\end{array}}{\mathcal{N};\mathcal{T};\Gamma \Downarrow^{\alpha+n}_{\rho} a\, t_1 \ldots t_n : \tau}$$

$$\textsc{Cut}\frac{\mathrm{ord}(\sigma \to \tau) \le \rho \quad \mathcal{N};\mathcal{T};\Gamma \Downarrow^{\alpha}_{\rho} s : \sigma \to \tau \quad \mathcal{N};\mathcal{T};\Gamma \Downarrow^{\alpha}_{\rho} t : \sigma}{\mathcal{N};\mathcal{T};\Gamma \Downarrow^{\alpha+1}_{\rho} s\, t : \tau}$$

There are a few interesting points to note. First, observe that the Var-, Term- and δ-rules all potentially "remove" more than one argument from the term, in turn increasing the height-bound α by more than 1. Also note that the δ-rule, as opposed to the Def-rule of the HORS typing system, does not replace the context Γ with the (left-over) argument variables but instead simply extends it. To make sure that t^a still does not use any variables outside its arguments, the premise $\mathcal{N};\mathcal{T};\mathrm{args}(\tau_1 \to \ldots \to \tau, \vec{x}) \vdash t^a : 0$ was added. Most of the design decisions of the derivation system are easily to motivate after the proof of Corollary 11, meaning we postpone the discussion of these to that point.

As the judgments of $\mathcal{N};\mathcal{T};\Gamma \Downarrow^{\alpha}_{\rho} t : \tau$ are quite big, we omit parts of them which are irrelevant or clear from the context wherever possible. Often, we simply write $\Downarrow^{\alpha}_{\rho} t$.

The remainder of this section is concerned with proving the validity of various transformations of derivations of $\Downarrow^{\alpha}_{\rho} t$. These culminate in Lemma 7 in which we show that each $\mathcal{N};\mathcal{T};\Gamma \vdash t : \tau$ may be embedded as a derivation of $\Downarrow^{\alpha}_{\rho} t$ in various ways.

Lemma 1 (Weakening) If $\Downarrow^{\alpha}_{\rho} s$, $\alpha < \alpha'$ and $\rho < \rho'$ then $\Downarrow^{\alpha'}_{\rho'} s$.

Lemma 2 (Renaming) If $\mathcal{N};\mathcal{T};\Gamma \Downarrow^{\alpha}_{\rho} s$ where $x : \tau, y : \tau \in \Gamma$ then $\mathcal{N};\mathcal{T};\Gamma \Downarrow^{\alpha}_{\rho} s[y/x]$.

Proof (Lemmas 1 and 2) Immediate by induction on $\mathcal{N};\mathcal{T};\Gamma \Downarrow^{\alpha}_{\rho} s$. ■

Lemma 3 (Appending) If $\mathcal{N};\mathcal{T};\Gamma \Downarrow^{\alpha}_{\rho} s : \sigma \to \tau$ and $y : \sigma \in \Gamma$, then $\mathcal{N};\mathcal{T};\Gamma \Downarrow^{\alpha+1}_{\rho} s\, y : \tau$.

Proof Begin by observing that $\Downarrow^{\alpha}_{\rho} y$ for any α via the Var rule and Lemma 1. We proceed via induction on $\Downarrow^{\alpha}_{\rho} s$.

Case Var: Then $s = x\,\vec{t}$ with $\Downarrow^{\alpha}_{\rho} t_i$ for $1 \leq i \leq n$, meaning $\Downarrow^{\alpha+n+1}_{\rho} x\,\vec{t}\,y$ as $\Downarrow^{\alpha}_{\rho} y$ as well.

Case γ: This case is impossible as $\mathbb{N};\mathcal{T};\Gamma \vdash s_0 + s_1 : 0$.

Case δ: Then $s = a\,t_1 \dots t_n$ with $\Downarrow^{\alpha}_{\rho} t^a[\vec{t}/\vec{x}]$, meaning $\Downarrow^{\alpha}_{\rho} t^a[\vec{t}/\vec{x}][y/x_{n+1}]$ via Lemma 2. Together with $\Downarrow^{\alpha}_{\rho} t_i$ and $\Downarrow^{\alpha}_{\rho} y$ this yields $\Downarrow^{\alpha+n+1}_{\rho} a\,t_1 \dots t_n\, y$.

Case Cut: Then $s = t\,t'$ and $\mathrm{ord}(t) \leq \rho$. But this means that $\mathrm{ord}(t\,t') \leq \mathrm{ord}(t) \leq \rho$ as well. We can thus arrive at $\Downarrow^{\alpha+1}_{\rho} t\,t'\,y$ via another Cut. ■

Lemma 4 (Substitution) If $\mathcal{N};\mathcal{T};\Gamma \Downarrow^{\alpha}_{\rho} s$ and $\mathcal{N};\mathcal{T};\Gamma \Downarrow^{\beta}_{\rho} t_i : \tau_i$ where $\mathrm{ord}(\tau_i) \leq \rho$ and $x_i : \tau_i \in \Gamma$ for $1 \leq i \leq n$ then $\mathcal{N};\mathcal{T};\Gamma \Downarrow^{\alpha+\beta}_{\rho} s[\vec{t}/\vec{x}]$.

Proof Proof per induction on $\Downarrow^{\alpha}_{\rho} s$. We write s^* as shorthand for $s[\vec{t}/\vec{x}]$

Case Var: Then $s = y\,\vec{u}$. Per inductive hypothesis we immediately obtain $\Downarrow^{\alpha+\beta}_{\rho} u_i^*$. If $y \notin \vec{x}$ then $y^* = y$ and we may thus conclude that $\Downarrow^{\alpha+n+\beta}_{\rho} y\,\vec{u^*}$. If $y \in \vec{x}$ then $y^* = t_j$. We know that $\mathrm{ord}(t_j) \leq \rho$ meaning we may derive $\Downarrow^{\alpha+n+\beta}_{\rho} t_j\,\vec{u^*}$ by an n-fold application of the Cut rule.

Case γ: Then $s = s_0 + s_1$. Per inductive hypothesis we obtain $\Downarrow^{\alpha+\beta}_{\rho} s_i^*$ leading us to conclude $\Downarrow^{\alpha+1+\beta}_{\rho} s_0^* + s_1^*$.

Case δ: Then $s = a\,\vec{u}$ with $(a : \tau \coloneqq \vec{y}, \vec{y'} \mapsto t^a) \in \mathcal{N}$. We know by inductive hypothesis that $\Downarrow^{\alpha+\beta}_{\rho} u_i^*$ and $\Downarrow^{\alpha+\beta}_{\rho} t^a[\vec{u}/\vec{y}][\vec{t}/\vec{x}]$. As $\mathcal{N};\mathcal{T};\mathrm{args}((\vec{y}, \vec{y'}), \tau) \vdash t^a : 0$ we know that $\vec{x} \notin \mathrm{FV}(t^a)$ and thus that $t^a[\vec{t}/\vec{x}] = t^a$. Then $t^a[\vec{u}/\vec{y}][\vec{t}/\vec{x}] = t^a[\vec{u^*}/\vec{y}]$ which yields $\Downarrow^{\alpha+\beta}_{\rho} t^a[\vec{u^*}/\vec{y}]$, allowing us to conclude $\Downarrow^{\alpha+n+\beta}_{\rho} a\,\vec{u^*}$ as desired.

Case Cut: Then $s = t_0\,t_1$. Per inductive hypothesis we know $\Downarrow^{\alpha+\beta}_{\rho} t_i^*$. As furthermore $\mathrm{ord}(t_0) = \mathrm{ord}(t_0^*)$ we conclude $\Downarrow^{\alpha+1+\beta}_{\rho} t_0^*\,t_1^*$. ■

With these transformations, we can prove the Cut-admissibility Lemma. It allows us to perform Cuts of order $\rho + 1$ in a derivation of $\Downarrow^{\alpha}_{\rho}$ "at the cost" of a more drastic increase of the height-bound of the resulting derivation.

Lemma 5 (Cut Admissibility) If $\mathcal{N};\mathcal{T};\Gamma \Downarrow^{\alpha}_{\rho} s : \sigma \to \tau$ where $\mathrm{ord}(\sigma \to \tau) \leq \rho + 1$ and $\mathcal{N};\mathcal{T};\Gamma \Downarrow^{\beta}_{\rho} t : \sigma$ then $\mathcal{N};\mathcal{T};\Gamma \Downarrow^{\alpha+\beta+1}_{\rho} s\,t : \tau$.

Proof Pick some $x \notin \Gamma$ then by Lemma 3 we have $\mathcal{N};\mathcal{T};\Gamma,(x:\sigma) \Downarrow^{\alpha+1}_{\rho} s\,x : \tau$. We may then apply Lemma 4 to obtain $\mathcal{N};\mathcal{T};\Gamma \Downarrow^{\alpha+\beta+1}_{\rho} s\,t : \tau$. ∎

We can then replace each Cut-application of the highest order with an application of the Cut-admissibility Lemma to lower the Cut-level ρ at the cost of an exponential blowup of the height-bound α.

Lemma 6 (Cut Reduction) If $\Downarrow^{\alpha}_{\rho+1} s$ then $\Downarrow^{2^{\alpha}}_{\rho} s$.

Proof We show $\Downarrow^{2^{\alpha}-1}_{\rho} s$ per induction on $\Downarrow^{\alpha}_{\rho}$. The claim follows by Lemma 1.

Case Var: Then $s = x\,\vec{t}$. We directly obtain $\Downarrow^{2^{\alpha}+n-1}_{\rho} x\,\vec{t}$ as $\Downarrow^{2^{\alpha}-1}_{\rho} t_i$ per inductive hypothesis. As $2^{\alpha}+n-1 \leq 2^{\alpha+n}-1$ the claim then follows by Lemma 1.

Case γ: Then $s = s_0 + s_1$. The inductive hypothesis yields $\Downarrow^{2^{\alpha}-1}_{\rho} s_i$. We then derive $\Downarrow^{2^{\alpha}}_{\rho} s_0 + s_1$ which we can weaken to $\Downarrow^{2^{\alpha+1}-1} s_0 + s_1$ via Lemma 1.

Case δ: Then $s = a\,\vec{t}$. Analogously to the Var case, we obtain $\Downarrow^{2^{\alpha}+n-1}_{\rho} a\,\vec{t}$ and apply Lemma 1 to derive the claim.

Case Cut: Then $s = t\,t'$. The inductive hypothesis yields $\Downarrow^{2^{\alpha}-1}_{\rho} t$ and $\Downarrow^{2^{\alpha}-1}_{\rho} t'$. Then an application of Lemma 5 yields precisely $\Downarrow^{2^{\alpha+1}-1}_{\rho} t\,t'$. ∎

The last result of this section shows that all well-typed acyclic HORS terminate by embedding $\mathcal{N};\mathcal{T};\Gamma \vdash t : \tau$ into $\Downarrow$. There are two different methods of achieving this. The "naïve" method simply replaces each $\vdash$-rule with its corresponding $\Downarrow$-rule. Notably, this means App is replaced with Cut, resulting in a height-bound of $h(t)$ and a Cut-level of $\deg(t)$. If we instead opt to replace each App of order $\deg(t)$ with an application of Lemma 6, we lower the Cut-level of the resulting derivation to $\deg(t)-1$ while increasing the height-bound to $s(t)$. As $s(t) \leq 2^{h(t)}$ the bound obtained via this embedding will usually be "more efficient" as one Cut-reduction step and thus one exponential blowup of α is prevented.

Lemma 7 (Embedding) Let $\mathcal{N};\mathcal{T};\Gamma \vdash t : \tau$. Then

1. If $\mathrm{grd}(t) \leq \rho$ then $\mathcal{N};\mathcal{T};\Gamma \Downarrow^{h(t)}_{\rho} t : \tau$.
2. If $\mathrm{grd}(t) \leq \rho + 1$ then $\Downarrow^{s(t)}_{\rho} t$.

Proof We prove $\Downarrow^{s(t)-1}_{\rho} t$ and $\Downarrow^{h(t)}_{\rho} t$ per induction on $\Gamma;\mathcal{N} \vdash t : \tau$. We combine the handling of both claims when possible.

Case Var: Then $t = x$ and thus $\Downarrow^{0}_{\rho} x$ by the Var rule.

Case Term: Then $t = a$ with $a \in \mathcal{T}$ and thus $\Downarrow_\rho^0 a$ by the Term rule.

Case Sum: Then $t = t_0 + t_1$. Per inductive hypothesis we have $\Downarrow_\rho^{\alpha_0} t_0$ and $\Downarrow_\rho^{\alpha_1} t_1$ yielding $\Downarrow_\rho^{\max\{\alpha_0,\alpha_1\}+1} t_0 + t_1$ which is sufficient, as for $\alpha_i = s(t_i) - 1$ we have

$$s(t_0 + t_1) - 1 = s(t_0) + s(t_1) \geq \max\{s(t_0), s(t_1)\} = \max\{s(t_0) - 1, s(t_1) - 1\} + 1$$

and for $\alpha_i = h(t_i)$ we have $h(t_0 + t_1) = \max\{h(t_0), h(t_1)\} + 1$.

Case Def: Then $t = a$ with $(a := \vec{x} \mapsto t^a) \in \mathcal{N}$. Per inductive hypothesis we obtain $\Downarrow_\rho^\alpha t^a$ and thus $\Downarrow_\rho^{\alpha+1} a$. This is the desired claim as $h(a) = h(t^a) + 1$ and $s(a) = s(t^a) + 1$.

Case App: Then $t = u\,u'$.

1. Per inductive hypothesis we obtain $\Downarrow_\rho^{h(u)} u$ and $\Downarrow_\rho^{h(u')} u'$. As we know $\mathrm{ord}(u) \leq \rho$ we obtain $\Downarrow_\rho^{h(u\,u)} u\,u'$ via Cut as $h(u\,u') = \max\{h(u), h(u')\} + 1$.
2. The inductive hypothesis yields $\Downarrow_\rho^{s(u)-1} u$ and $\Downarrow_\rho^{s(u')-1} u'$. Because $\mathrm{ord}(u) \leq \rho + 1$ we can apply Lemma 5 to obtain $\Downarrow_\rho^{s(u\,u)-1} u\,u'$ as $s(u\,u') = s(u) + s(u') + 1$. ∎

4 Reduction Chain Length Bounds

4.1 Deriving an Upper Bound

This section is concerned with extracting an upper bound on the length of derivation chains for HORSes. The measures we want to approximate are

$$\begin{aligned} d(\mathcal{N}; t) &:= \max\{n < \omega \mid \text{there is a chain } \mathcal{N} \vdash t_0 \rightsquigarrow t_1 \rightsquigarrow \ldots \rightsquigarrow t_n \text{ with } t_0 = t\} \\ ds_n(N) &:= \max\{d(\mathcal{N}; t) \mid \Gamma, \mathcal{T}, \mathcal{N} \vdash t : \tau \text{ and } \deg(t) \leq n \text{ and } s(t) \leq N\} \\ dh_n(N) &:= \max\{d(\mathcal{N}; t) \mid \Gamma, \mathcal{T}, \mathcal{N} \vdash t : \tau \text{ and } \deg(t) \leq n \text{ and } h(t) \leq N\} \end{aligned}$$

As $\mathcal{N}$ is usually clear from the context, we simply write $d(t)$ from now on. Observe also that $ds_n(N)$ and $dh_n(N)$ can only take into account well-typed t as these are the only ones for which $\deg(t)$ can be determined.

We now define a size measure on Cut-free derivations. Observe that as the structure of a derivation $\Downarrow_0^\alpha t$ is completely determined by the syntactical structure of t (and the contents of $\mathcal{N}$) it is reasonable to write $|t|$ as it is thus uniquely defined.

Definition 8 For $\Downarrow_0^\alpha t$ we define **its size $|t|$** recursively on $\Downarrow_0^\alpha t$:

$$\begin{aligned} |x\, t_1 \ldots t_n| &:= \Sigma_{i=1}^n |t_i| && (\text{Var}) \\ |a\, t_1 \ldots t_n| &:= \Sigma_{i=1}^n |t_i| && (\text{Term}) \\ |s_0 + s_1| &:= 1 + |s_0| + |s_1| && (\gamma) \\ |a\, t_1 \ldots t_n| &:= 1 + |t^a[\vec{t}/\vec{x}]| + \Sigma_{i=1}^n |t_i| && (\delta) \end{aligned}$$

Lemma 9 (Estimation) If $\Downarrow_0^\alpha s$ then $|s| \leq 2^\alpha$.

Proof We prove $|s| \leq 2^\alpha - 1$ per induction on $\Downarrow_\rho^\alpha s$.

Case VAR: Per inductive hypothesis $|t_i| \leq 2^\alpha - 1$. Then

$$|x\,\vec{t}| = \Sigma_{i=1}^n |t_i| \leq \Sigma_{i=1}^n 2^\alpha - 1 = n(2^\alpha - 1) \leq 2^{\alpha+n} - 1$$

Case γ: Per inductive hypothesis $|s_i| \leq 2^\alpha - 1$. Then

$$|s_0 + s_1| = 1 + |s_0| + |s_1| \leq 1 + 2(2^\alpha - 1) = 2^{\alpha+1} - 1$$

Case δ: Per inductive hypothesis $|t^a[\vec{t}/\vec{x}]|, |t_i| \leq 2^\alpha - 1$. Then

$$|a\,\vec{t}| = 1 + |t^a[\vec{t}/\vec{x}]| + \Sigma_{i=1}^n |t_i| \leq 1 + (n+1)(2^\alpha - 1) \leq 2^{\alpha+n+1} - 1$$ ∎

The following lemma is crucial to the extracting the upper bounds as it connects the expanded head reduction trees with reduction chains. Indeed, this lemma directly yields a bound on reduction chains (Corollary 11) and a proof of the strong normalization of acyclic HORS (Corollary 12).

Lemma 10 If $\Downarrow_0^\alpha s$ and $s \rightsquigarrow^+ t$ then $\Downarrow_0^\beta t$ for some β and $|s| > |t|$.

Proof Observe that the claim can be extended to the case $s \rightsquigarrow^* t$ weakening the inequality to $|s| \geq |t|$. We prove the claim per induction on $\Downarrow_\rho^\alpha s$.

Case VAR: Then $s = x\,\vec{u}$. We know that $x\,\vec{u} \rightsquigarrow^+ t$ means that $x\,u_1 \dots u_n \rightsquigarrow^+ x\,u_1' \dots u_n'$ where $u_i \rightsquigarrow^* u_i'$. Per inductive hypothesis, we know that $\Downarrow_\rho^{\beta_i} u_i'$, meaning if we take $\beta := \max\{\beta_1, \dots, \beta_n\}$ we have $\Downarrow_\rho^\beta u_i$ for $1 \leq i \leq n$. Thus $\Downarrow_\rho^{\beta+n} x\,\vec{u}'$. Per inductive hypothesis, we know $|u_i| \geq |u_i'|$. Further more, $x\,\vec{u} \rightsquigarrow^+ x\,\vec{u}'$ means there is $u_j \rightsquigarrow^+ u_j'$ for which we have $|u_j| > |u_j'|$. This yields the desired inequality:

$$|x\,\vec{u}| = \sum_{i=1}^n |u_i| > \sum_{i=1}^n |u_i'| = |x\,\vec{u}'|$$

Case γ: Then $s = s_0 + s_1$. Again, there are two possibilities how $s_0 + s_1 \rightsquigarrow^+ t'$ can be constituted.

1. Possibly, $s_0 + s_1 \rightsquigarrow^+ s_0' + s_1'$ where $s_i \rightsquigarrow^* s_i'$. Then we obtain $s_i \rightsquigarrow^* s_i'$ where at least for one $i \in \{0, 1\}$ we even have $s_i \rightsquigarrow^+ s_i'$. Per inductive hypothesis this yields $\Downarrow_0^{\beta_i} s_i'$ and thus $\Downarrow_0^{\beta+1} s_0' + s_1'$ if we take $\beta := \max\{\beta_0, \beta_1\}$. Because $|s_i| > |s_i'|$ for at least one $i \in \{0, 1\}$ we also have

$$|s_0 + s_1| = 1 + |s_0| + |s_0| > 1 + |s_0'| + |s_0'| = |s_0' + s_1'|$$

2. Otherwise, $s_0 + s_1 \rightsquigarrow^* s_0' + s_1' \rightsquigarrow s_i' \rightsquigarrow^* t'$ where $s_i \rightsquigarrow^* s_i'$. Then we obtain $s_i \rightsquigarrow^* s_i' \rightsquigarrow^* t'$ by replaying the $s_i \rightsquigarrow^* s_i'$. Per inductive hypothesis, this yields $\Downarrow_0^\beta t'$ and $|s_i| \geq |t'|$. Then we can conclude

$$|s_0 + s_1| = 1 + |s_0| + |s_1| > |s_i| \geq |t'|$$

Case δ**:** Then $s = a\,\vec{u}$. There are two possibilities how $a\,\vec{u} \rightsquigarrow^+ t$ can be constituted.

1. Possibly, $a\,u_1 \ldots u_n \rightsquigarrow^+ a\,u_1' \ldots u_n'$ with $u_i \rightsquigarrow^* u_i'$ and $u_j \rightsquigarrow^+ u_j'$ for at least one j. Then we can obtain $t^a[\vec{u}/\vec{x}] \rightsquigarrow^* t^a[\vec{u'}/\vec{x}]$ by simply replaying $u_i \rightsquigarrow^* u_i'$ at each substituted instance of u_i. This yields $\Downarrow_0^{\beta_0}$ and $|t^a[\vec{u}/\vec{x}] \geq |t^a[\vec{u'}/\vec{x}]||$ per inductive hypothesis. For each u_i, we obtain $\Downarrow_0^{\beta_i} u_i'$ and $|u_i| \geq |u_i'|$ per inductive hypothesis, with $|u_j| > |u_j'|$. Then clearly $\Downarrow_0^{\beta+n+1} a\,\vec{u'}$ with $\beta = \max\{\beta_0, \ldots, \beta_n\}$. Furthermore

$$|a\,\vec{u}| = 1 + |t^a[\vec{u}/\vec{x}]| + \sum_{i=1}^{n} |u_i| > 1 + |t^a[\vec{u'}/\vec{x}]| + \sum_{i=1}^{n} |u_i'| = |a\,\vec{u'}|$$

2. Otherwise, $a\,u_1 \ldots u_n \rightsquigarrow^* a\,u_1' \ldots u_n' \rightsquigarrow t^a[\vec{u'}/\vec{x}] \rightsquigarrow^* t'$ with $u_i \rightsquigarrow^* u_i'$. Then we can derive $t^a[\vec{u}/\vec{x}] \rightsquigarrow^* t^a[\vec{u'}/\vec{x}]$ as in the case above. Together with $t^a[\vec{u'}/\vec{x}] \rightsquigarrow^* t'$ this yields $t^a[\vec{u}/\vec{x}] \rightsquigarrow^* t'$ which allows us to derive $\Downarrow_0^\beta t'$ and $|t^a[\vec{u}/\vec{x}]| \geq |t'|$ per inductive hypothesis on $t^a[\vec{u}/\vec{x}]$. Then

$$|a\,\vec{u}| = 1 + |t^a[\vec{u}/\vec{x}]| + \sum_{i=1}^{n} |t_i| > |t^a[\vec{u}/\vec{x}]| \geq |t'|$$

∎

With Lemma 10 in mind, we can discuss the motivations behind the design of the $\Downarrow_\beta^\alpha$ derivation system. A simplified system with all irrelevant details omitted is given below.

$$\textsc{Var}\,\frac{\Downarrow_\rho^\alpha t_1 \quad \ldots \quad \Downarrow_\rho^\alpha t_n}{\Downarrow_\rho^{\alpha+n} x\,\vec{t}} \qquad \delta\,\frac{\Downarrow_\rho^\alpha t^a[\vec{t}/\vec{x}] \quad \Downarrow_\rho^\alpha t_1 \quad \ldots \quad \Downarrow_\rho^\alpha t_n}{\Downarrow_\rho^{\alpha+n+1} a\,\vec{t}}$$

$$+\,\frac{\Downarrow_\rho^\alpha s_0\,\vec{t} \quad \Downarrow_\rho^\alpha s_1\,\vec{t}}{\Downarrow_\rho^{\alpha+1} (s_0 + s_1)\,\vec{t}}$$

Lemma 10 essentially relies on the fact that whenever a term $\Downarrow_0^\alpha s$ takes a reduction step $s \rightsquigarrow t$, this reduction has to be “accounted for” somewhere within the derivation of $\Downarrow_0^\alpha s$. Thus each of the (non-Cut) rules has as its premises the various “places” a reduction step could lead to or take place in. Simplest of all, the Var- and Term-rules only have to account for their argument terms as variables and terminals cannot take reduction steps on their own, so any reduction involving them will have taken place in one of their arguments. For the case of sums, we of course need to account for the two reductions $s_0 + s_1 \rightsquigarrow s_i$.

Notably, this already suffices to also account for reductions in the arguments to the sum as these can just be "replayed" after either sum reduction step has taken place. The most interesting case is that of δ: Clearly, we need to account for $\Downarrow_0^{\alpha}\ t^a[\vec{t}/\vec{x}]$ as this reduction step may be taken by the non-terminal itself. Differing from the case for sums, this does not suffice to also account for the non-terminal's arguments, as possibly $x_i \notin \mathrm{FV}(t^a)$, meaning $t_i \notin \mathrm{FV}(t^a[\vec{t}/\vec{x}])$ which would in turn make it impossible to "replay" reduction steps in t_i on $t^a[\vec{t}/\vec{x}]$. Thus the δ-rule needs to also account for all of the arguments t_i on their own as well.

Corollary 11 (Boundedness) For any $\mathcal{N};\mathcal{T};\Gamma \vdash s : \tau$ we have $d(s) \leq |s|$.

Proof For any (possibly infinite) reduction chain $s = s_0 \rightsquigarrow s_1 \rightsquigarrow s_2 \rightsquigarrow \ldots$ we know that $|s_i| > |s_{i+1}|$ by Lemma 10. That means any chain can be at most of length $|s|$. ■

Corollary 12 (Strong Normalization) All reduction chains of $\mathcal{N};\mathcal{T};\Gamma \vdash s : \tau$ are finite.

Combining all of the results from Section 3 and this section, we can explicitly derive the desired upper bounds.

Theorem 13 (Upper bounds)

$$dh_n(N) \leq 2_{n+1}(N) \qquad\qquad ds_n(N) \leq 2_n(N)$$

Proof Let $\mathcal{N};\mathcal{T};\Gamma \vdash t : \tau$ with $\deg(t) = n$ and $n > 0$. By Lemma 7 we know that

$$\mathcal{N};\mathcal{T};\Gamma \Downarrow_n^{h(t)}\ t : \tau \qquad\qquad \mathcal{N};\mathcal{T};\Gamma \Downarrow_{n-1}^{s(t)}\ t : \tau$$

Now we can repeatedly apply Lemma 6 to obtain

$$\mathcal{N};\mathcal{T};\Gamma \Downarrow_0^{2_n(h(t))}\ t : \tau \qquad\qquad \mathcal{N};\mathcal{T};\Gamma \Downarrow_0^{2_{n-1}(s(t))}\ t : \tau$$

which allows us to estimate, using Lemma 9, that

$$|t| \leq 2_{n+1}(h(t)) \qquad\qquad |t| \leq 2_n(s(t))$$

Lastly, because of Corollary 11 we know

$$d(t) \leq |t| \leq 2_{n+1}(h(t)) \qquad\qquad d(t) \leq |t| \leq 2_n(s(t))$$

As we can do this for any $\mathcal{N};\mathcal{T};\Gamma \vdash t : \tau$ with $\deg(t) = n$ this yields the bounds as desired.

For the case of $n = 0$, observe we must have $\mathcal{N};\mathcal{T};\Gamma \vdash t : 0$ where t must be a sum of variables, terminals and non-terminals of type 0, the non-terminals in turn unfolding to analogous terms of degree 0. For such terms, each step $t \rightsquigarrow t'$ guarantees $s(t) > s(t')$ and $h(t) > h(t')$ as no substitutions take place. Thus we obtain

$$dh_0(N) \leq N < 2_1(N) \qquad\qquad ds_0(N) \leq N = 2_0(N)$$

■

4.2 Deriving a Lower Bound

In this section, we show that the upper bounds from Section 4.1 are, while not necessarily fully exact, at least very reasonable approximations of the desired bounds. More precisely, we prove that

$$ds_n(N) = 2_n(\Omega(N)) \qquad dh_n(N) = 2_{n+1}(\Omega(N))$$

from which we can conclude that the bounds from Section 4.1 can only be improved to something of the shape

$$ds_n(N) \leq 2_n\left(\lfloor r_s N \rfloor + c_s\right) \qquad dh_n(N) \leq 2_{n+1}\left(\lfloor r_h N \rfloor + c_h\right)$$

for constants $0 < r_s, r_h < 1$ and $c_s, c_h \in \mathbb{R}$.

To derive this result, we define terms $D^s_n(N)$ and $D^h_n(N)$ of degree n and with $s(D^s_n(N)) = O(N)$ and $h(D^h_n(N)) = O(N)$ with $d(D^h_n(N)) > 2_n(N)$ and $d(D^h_n(N)) > 2_{n+1}(N)$. For this, let $b : 0 \to 0 \to 0$ be a terminal. For terms $t : 0$ we define b-trees of height K with t at each leaf as follows

$$T_0(t) \coloneqq t \qquad T_{K+1}(t) \coloneqq b\, T_K(t)\, T_K(t)$$

Now, with a non-terminal $B : 0 \to 0 \coloneqq x \mapsto b\,x\,x$ and writing $[f]^n(v)$ for the n-fold application of f to v as defined below

$$[f]^0(v) \coloneqq v \quad [f]^{n+1}(v) \coloneqq f([f]^n(v))$$

we get the following result.

Fact 14 For any $t : 0$ we have $[B]^K\, t \rightsquigarrow^* T_K(t)$

Proof Simple induction on K. ■

The remainder of this section is concerned with "compressing" $T_{2_{n-1}(N)}(t)$ into a term $s(S^N_n(t)) = O(N)$ with $S^N_n(t) \rightsquigarrow^* T_{2_{n-1}(N)}(t)$ and $T_{2_n(N)}(t)$ into a term $h(H^N_n(t)) = O(N)$ with $H^N_n(t) \rightsquigarrow^* T_{2_n(N)}(t)$.

In [1] this is done with the help of Church encodings of natural numbers, an approach that we parallel in HORS. Unfortunately, the reduction behavior of HORS Church encodings of different orders applied to each other is somewhat erratic, as one often may only take reduction steps in the outer-most application. We thus first develop an abstract theory of n-repeaters, which can be seen as a semantic characterization of Church encodings, which helps us control this behavior in our proofs. Writing τ_i for the types

$$\tau_0 \coloneqq 0 \qquad \tau_{i+1} \coloneqq \tau_i \to \tau_i$$

and $\mathrm{tower}(n_0, \ldots, n_m)$ for the tower of exponentials

$$\mathrm{tower}(n) \coloneqq n \qquad \mathrm{tower}(n, \vec{m}) \coloneqq n^{\mathrm{tower}(\vec{m})}$$

we inductively define n-repeaters below.

Definition 15

0. We call an expression $f_0^n : \tau_2$ an **n-repeater of order 0** if $n > 0$ and for any $m : \tau_1$ and $v : \tau_0$ we have $f_0^n\, m\, v \rightsquigarrow^* [m]^n(v)$

i. We call an expression $f_i^{n_i} : \tau_{i+2}$ an **n_i-repeater of order i** if $n_i > 0$ and for any family $(f_j^{n_j})_{j<i}$ of n_j-repeaters of order j and any $m : \tau_1$ and $v : \tau_0$ we have that $f_i^{n_i}\, f_{i-1}^{n_{i-1}} \ldots f_0^{n_0}\, m\, v \rightsquigarrow^* [m]^k(v)$ where $k = \mathrm{tower}(n_0, n_1, \ldots, n_i)$.

With this definition, we can prove a crucial Lemma. It is needed for proving both the well-definedness of our HORS Church encodings and the correctness of $S_n^N(t)$ and $H_n^N(t)$.

Lemma 16 For any family $(f_j^{n_j})_{j\leq i}$ of n_j-repeaters of order j, any $N > 0$ and any $f_{-1} : \tau_1$ and $f_{-2} : \tau_0$ we have $[f_i^{n_i}]^N(f_{i-1}^{n_{i-1}}) \ldots f_{-1}\, f_{-2} \rightsquigarrow^* [f_{-1}]^k(f_{-2})$ with $k = \mathrm{tower}(n_0, \ldots, n_i, N)$.

Proof For the sake of readability, we take $m \coloneqq f_{-1}$ and $v \coloneqq f_{-2}$ again. We proceed by induction on the maximum order i.

$i = 0:$ We proceed by induction on N.

$N = 1:$ Then $[f_0^n]^1(m)\, v = f_0^n\, m\, v \rightsquigarrow^* [m]^n(v) = [m]^{n^1}(v)$.

$N + 1:$

$$\begin{aligned} [f_0^n]^{N+1}(m)\, v = f_0^n\, ([f_0^n]^N(m))\, v &\rightsquigarrow^* [[f_0^n]^N(m)]^n(v) \\ &= \underbrace{[f_0^n]^N(m)(\ldots([f_0^n]^N(m)}_{n\text{ times}}\, v)\ldots) \\ &\overset{\mathrm{IH}}{\rightsquigarrow^*} \underbrace{[m]^{n^N}(\ldots([m]^{n^N}}_{n\text{ times}}(v))\ldots) = [m]^{n^{N+1}}(v) \end{aligned}$$

$i + 1:$ Again, we perform an induction on N.

$N = 1:$ Analogous to the case for $i = 0$.

$N + 1:$ Then $[f_i^{n_i}]^{N+1}(f_{i-1}^{n_{i-1}}) \ldots m\, v = f_i^{n_i}([f_i^{n_i}]^N(f_{i-1}^{n_{i-1}})) \ldots m\, v \rightsquigarrow^* [m]^k(v)$ with $k = \mathrm{tower}(n_0, \ldots, n_{i-2}, n_{i-1}^{n_i^N}, n_i)$ because, per inductive hypothesis, the expression $[f_i^{n_i}]^N(f_{i-1}^{n_{i-1}})$ is an $n_{i-1}^{n_i^N}$-repeater of order $i-1$. Now, as $(n_{i-1}^{n_i^N})^{n_i} = n_{i-1}^{n_i * n_i^N} = n_{i-1}^{n_i^{N+1}}$, we can conclude that $k = \mathrm{tower}(n_0, \ldots, n_{i-1}, n_i, N+1)$. ∎

We now define the HORS Church encoding $e_i^n : (\tau_i \to \tau_i) \to \tau_i \to \tau_i$ of n specialized to values of order i as follows

$$e_i^n : \tau_{i+2} \coloneqq f\, x_i\, \vec{x} \mapsto [f]^n(x_i)\, \vec{x}$$

where $\vec{x} = x_{i-1} \dots x_0$.

Lemma 17 The term $e_i^n : \tau_{i+2}$ is a n-repeater of order i.

Proof Proof per induction on i.

$i = 0 :$ $e_0^n\, m\, v = [m]^n(v)$ per definition.

$i + 1 :$ If $(f_j^{n_j})_{j<i}$ are n_j-repeaters of order j then

$$e_i^n\, f_{i-1}^{n_{i-1}}\, f_{i-2}^{n_{i-2}} \dots m\, v \rightsquigarrow [f^{n_{i-1}}]^n(f_{i-2}^{n_{i-2}}) \dots m\, v \rightsquigarrow^* [m]^k(v)$$

with $k = \text{tower}(n_0, \dots, n_{i-1}, n)$ by Lemma 16. Observe that for $i = 1$ we have $f_{i-2}^{n_{i-2}} = m$ which does not cause any issues. ∎

Now we can define $S_n^N(t) \coloneqq [e_{n-2}^2]^N(e_{n-3}^2)\, e_{n-4}^2 \dots e_{-1}^2\, e_{-2}^2$, taking $e_{-1}^2 \coloneqq B$ and $e_{-2}^2 \coloneqq t$.

Fact 18

1. We have $S_n^N(t) \rightsquigarrow^* T_{2_{n-1}(N)}(t)$
2. For a fixed n, we have $s(S_n^N(t)) = O(N)$
3. We have $\deg(S_n^N(t)) = n$

Proof

1. As the e_i^2 are all n-repeaters of order i by Lemma 17, we know by Lemma 16 that

$$S_n^N(t) = [e_{n-2}^2]^N\,(e_{n-3}^2) \dots e_{-1}^2\, e_{-2}^2 \rightsquigarrow^* [e_{-1}^2]^{2_{n-1}(N)}(e_{-2}^2) = [B]^{2_{n-1}(N)}(t)$$

   and furthermore by Fact 14 that $[B]^{2_{n-1}(N)}(t) \rightsquigarrow^* T_{2_{n-1}(N)}(t)$.
2. Observe that for arbitrary m and v we have $s([m]^N(v)) = s(v) + N(s(m) + 1)$ and thus that $s(e_i^n) = n + i + 1$ and $s([e_{n-2}^2]^N(e_{n-3}^2)) = n + N(n + 1) = O(N)$. Together, this yields

$$s(S_n^N(t)) = \underbrace{s(t) + s(B) + 2 + \sum_{j=0}^{j<n-3} (s(e_j^2) + 1)}_{\text{constant for fixed } n} + \underbrace{s([e_{n-2}^2]^N(e_{n-3}^2))}_{O(N)} = O(N)$$

3. Observe that $\deg(S_n^N(t)) = \text{ord}(e_{n-2}^2) = n$ as $e_i^2 : \tau_{i+2}$. ∎

Unfortunately, this approach does not extend directly to the definition of $H_n^N(t)$. If we were to take $H_n^N(t) := S_n^{2^N}(t)$ this would yield the desired $d(H_N^n) \rightsquigarrow^* T_{2_n(N)}(t)$ however, we would also have $h(H_N^n) = O(2^N)$. We thus have to take an additional "compression step", squeezing $[e_{n-2}^2]^{2^N}$ into a term $h(c_i^N) = O(N)$. For this, we essentially take advantage of the fact a binary tree of height N may have 2^N nodes.

$$c_i^0 : \tau_{i+2} := f\,\vec{v} \mapsto f\,\vec{v} \qquad\qquad c_i^{N+1} : \tau_{i+2} := f\,\vec{v} \mapsto [c_i^N]^2(f)\,\vec{v}$$

Lemma 19 The term $c_i^N : \tau_{i+2}$ is a 2^N-repeater of order i.

Proof Denote by $(f_j^{n_j})_{j<i}$ a family n_j-repeaters of order j. Proof per induction on N.

$N = 0$: Then $c_i^0\, f_{i-1}^{n_{i-1}} \dots m\, v \rightsquigarrow f_{i-1}^{n_{i-1}} \dots m\, v \rightsquigarrow^* [m]^k(v)$ with $k = \text{tower}(n_0, \dots, n_{i-1}) = \text{tower}(n_0, \dots, n_{i-1}, 2^0)$.

$N + 1$: Then

$$c_i^{N+1}\, f_{i-1}^{n_{i-1}} \dots m\, v \rightsquigarrow [c_i^N]^2\, f_{i-1}^{n_{i-1}}\, m\, v \rightsquigarrow [m]^k(v)$$

with $k = \text{tower}(n_0, \dots, n_{i-1}, 2^N, 2)$ by Lemma 16 and the inductive hypothesis. As $(2^N)^2 = 2^{N+1}$ we then obtain $k = \text{tower}(n_0, \dots, n_{i-1}, 2^{N+1})$ as desired. ∎

Now we may simply take $H_n^N(t) := c_{n-2}^N\, e_{n-3}^2 \dots e_{-1}\, e_{-2}$ with $e_{-1} := B$ and $e_{-2} := t$.

Lemma 20

1. We have $H_n^N(t) \rightsquigarrow^* T_{2_n(N)}(t)$
2. For a fixed n, we have $h(H_N^n(t)) = O(N)$
3. We have $\deg(H_N^n(t)) = n$

Proof

1. We obtain
$$H_n^N(t) = c_{n-2}^N\, e_{n-3}^2 \dots B\,\text{on} \rightsquigarrow^* [B]^{2_n(N)}(t) \rightsquigarrow^* T_{2_n(N)}(t)$$
by Lemma 19 and Fact 14.
2. Inductively we can see $h(c_i^N) \le N(i+1) = O(N)$ and thus
$$h(H_n^N(t)) = \underbrace{h(c_i^N)}_{O(N)} + \underbrace{h(x\, e_{n-2}^2 \dots B\,\text{on})}_{\text{constant for fixed } n} = O(N)$$
3. Simply observe that $\deg(H_n^N(t)) = \text{ord}(c_{n-2}^N) = n$. ∎

To derive the desired bounds, we need one more result.

Lemma 21 For any $t, s : 0$ with $t \rightsquigarrow^n s$ we have $T_K(t) \rightsquigarrow^{n*2^K} T_K(s)$

Proof Observe that $T_K(t)$ has 2^K copies of t as its nodes. Then each copy can take n steps, leading to $n * 2^K$ steps from $T_K(t)$ to $T_K(s)$ overall. ∎

Now fix some terminal off and non-terminal on with $(\text{on} \mapsto \text{off}) \in \mathbb{N}$. Then clearly $D^s_n(N) := S^N_n(\text{on})$ and $D^h_n(N) := H^N_n(\text{on})$ are AHORS terms with the desired properties.

5 Abstract Rewriting Properties of AHORS

We begin by recalling some basic definitions and results about abstract rewriting systems.

Definition 22 Let $\rightsquigarrow$ be a binary relation on some set T of terms.

1. We call $\rightsquigarrow$ **terminating** if it is well-founded
2. We call $\rightsquigarrow$ **confluent** if for any $s, t, u \in T$ with $s \rightsquigarrow^* t$ and $s \rightsquigarrow^* u$ then there exists $v \in T$ with $t \rightsquigarrow^* v$ and $u \rightsquigarrow^* v$
3. We call $\rightsquigarrow$ **locally confluent** if for any $s, t, u \in T$ with $s \rightsquigarrow t$ and $s \rightsquigarrow u$ then there exists $v \in T$ with $t \rightsquigarrow^* v$ and $u \rightsquigarrow^* v$
4. We call $s \in T$ **normal** if s is $\rightsquigarrow$-maximal
5. We call $t \in T$ a **normal form** of $s \in T$ if t is normal and $s \rightsquigarrow^* t$

Lemma 23 If $\rightsquigarrow$ is confluent, any $s \in T$ has at most one normal form.

Proof Suppose t_0, t_1 were normal forms of s. As $s \rightsquigarrow^* t_i$ we know by the confluence of $\rightsquigarrow$ that there is some u with $t_i \rightsquigarrow^* u$. But because the t_i are normal, we know that $t_i \rightsquigarrow^* u$ have to be $\rightsquigarrow$-sequences of length 0 and we can conclude that $t_0 = u = t_1$. ∎

Lemma 24 (Newman) If $\rightsquigarrow$ is terminating and locally confluent, $\rightsquigarrow$ is confluent.

Proof We prove $\forall u, v.\ s \rightsquigarrow^* u \wedge s \rightsquigarrow^* v \rightarrow \exists t.\ u \rightsquigarrow^* t \wedge v \rightsquigarrow^* t$ by $\rightsquigarrow$-induction on s. Thus suppose that $s \rightsquigarrow^* u_0$ and $s \rightsquigarrow^* u_1$. The only non-trivial case is if there are t_0, t_1 such that $s \rightsquigarrow t_i \rightsquigarrow^* u_i$. Then

1. As $s \rightsquigarrow t_i$, local confluence yields a v such that $t_i \rightsquigarrow^* v$
2. Because $s \rightsquigarrow t_0$ we may apply the inductive hypothesis to $t_0 \rightsquigarrow^* u_0$ and $t_0 \rightsquigarrow^* v$ to obtain an x such that $u_0 \rightsquigarrow^* x$ and $v \rightsquigarrow^* x$
3. Again, as $s \rightsquigarrow t_1$, we may apply the inductive hypothesis to $t_1 \rightsquigarrow^* v \rightsquigarrow^* x$ and $t_i \rightsquigarrow^* u_1$ to obtain a y such that $x \rightsquigarrow^* y$ and $u_1 \rightsquigarrow^* y$

Then this y is the joining point with $u_0 \rightsquigarrow^* x \rightsquigarrow y$ and $u_1 \rightsquigarrow^* y$. ∎

In the analysis of the abstract rewriting properties of AHORS, it is useful to separate the transition relation into $\rightsquigarrow_\delta$ and $\rightsquigarrow_\gamma$, each only consisting of transition steps justified by the δ- and γ-rule, respectively. These two relations have very different abstract rewriting properties. For example, $\rightsquigarrow_\gamma$ is not even locally confluent: Consider the expression $s := 0 \mid 1$ where $0, 1 \in \mathcal{T}$, clearly $s \rightsquigarrow_\gamma 0$ and $s \rightsquigarrow_\gamma 1$ but as 0 and 1 are normal, there can be no joining point. On the other hand, $\rightsquigarrow_\delta$ is confluent.

Theorem 25 The transition relation $\rightsquigarrow_\delta$ is confluent.

Proof By Lemma 24 it suffices to show that $\rightsquigarrow_\delta$ is terminating and locally confluent. For termination, observe that the bounds from Theorem 13 also bound the length of $\rightsquigarrow_\delta$-reduction sequences, making $\rightsquigarrow_\delta$ well-founded. It remains to prove local confluence. Thus let $s \rightsquigarrow_\delta t$ and $s \rightsquigarrow_\delta u$. We prove the existence of some v with $t \rightsquigarrow_\delta^* v$ and $u \rightsquigarrow_\delta^* v$ per induction on $s \rightsquigarrow_\delta t$.

Sum: Then $s = x_0 \mid x_1$ and there are two possibilities how $s \rightsquigarrow_\delta t$ and $s \rightsquigarrow_\delta u$ are constituted. Either, wlog., $s \rightsquigarrow_\delta t$ is $x_0 \mid x_1 \rightsquigarrow_\delta x_0' \mid x_1$ and $s \rightsquigarrow_\delta t$ is $x_0 \mid x_1 \rightsquigarrow_\delta x_0 \mid x_1'$, meaning their joining point is simply $x_0' \mid x_1'$. Otherwise, and wlog, the transitions are $x_0 \mid x_1 \rightsquigarrow_\delta x_t \mid x_1$ and $x_0 \mid x_1 \rightsquigarrow_\delta x_u \mid x_1$. Then the inductive hypothesis yields a joining point x_0' of $x_0 \rightsquigarrow x_t$ and $x_0 \rightsquigarrow x_u$ such that $x_t \rightsquigarrow_\delta^* x_0'$ and $x_u \rightsquigarrow_\delta^* x_0'$. Then $x_0' \mid x_1$ is the joining point of $s \rightsquigarrow_\delta t$ and $s \rightsquigarrow_\delta u$.

App: Then $s = x_0\, x_1$ and we assume that $s \rightsquigarrow_\delta u$ is no δ-step, as this is handled, by symmetry, in the δ case. Then this case is analogous to the Sum case as we only need to distinguish whether $s \rightsquigarrow_\delta t$ and $s \rightsquigarrow_\delta u$ take place in the same x_i or not.

δ: Then $s = a\, x_0 \ldots x_n$ and $s \rightsquigarrow_\delta t$ is $a\, y_0 \ldots y_n \rightsquigarrow_\delta t_a[y_0/x_0, \ldots, y_n/x_n]$. If $s \rightsquigarrow_\delta u$ is the same, the claim is trivial. Thus, let $s \rightsquigarrow_\delta u$ instead boil down to some $y_i \rightsquigarrow_\delta y_i'$ and let, wlog, $i = 0$, meaning it is $a\, y_0 \ldots y_n \rightsquigarrow_\delta a\, y_0' \ldots y_n$. Then the joining point is $t_a[y_0'/x_0, \ldots, y_n/x_n]$ which is arrived at via $a\, y_0' \ldots y_n \rightsquigarrow_\delta t_a[y_0'/x_0, \ldots, y_n/x_n]$ and $t_a[y_0/x_0, \ldots, y_n/x_n] \rightsquigarrow_\delta^* t_a[y_0'/x_0, \ldots, y_n/x_n]$ which is obtained by taking $y_0 \rightsquigarrow_\delta y_0'$ at every substituted-in instance of y_0. ∎

Corollary 26 Every AHORS has a unique δ-normal form.

The main result we aim to obtain in this section is about the interaction between γ- and δ-transitions.

Lemma 27 Let $s \rightsquigarrow_\gamma t \rightsquigarrow_\delta u$ then there exists a t' such that $s \rightsquigarrow_\delta t' \rightsquigarrow_\gamma^* u$.

Proof As in Theorem 25, this is trivial if the transitions take place in separate subterms of s, for example if $s = x_0 \mid x_1$ and $x_0 \rightsquigarrow_\gamma x_0'$, $x_1 \rightsquigarrow_\delta x_1'$. There are only two tricky cases to consider:

1. If $s = a\,y_0 \ldots y_n$ and wlog. $a\,y_0 \ldots y_n \rightsquigarrow_\gamma a\,y_0' \ldots y_n \rightsquigarrow_\delta t_a[y_0'/x_0, \ldots, y_n/x_n]$. Then choose $t' = t_a[y_0/x_0, \ldots, y_n/x_n]$ with the sequence $a\,y_0 \ldots y_n \rightsquigarrow_\delta t_a[y_0/x_0, \ldots, y_n/x_n] \rightsquigarrow_\gamma^* t_a[y_0'/x_0, \ldots, y_n/x_n]$ where the last part is obtained by taking $y_0 \rightsquigarrow_\gamma y_0'$ at all instances of y_0.
2. If $s = x_0 \mid x_1$ and wlog. $x_0 \mid x_1 \rightsquigarrow_\gamma x_0 \rightsquigarrow_\delta x_0'$. Then choose $t' = x_0' \mid x_1$ with $x_0 \mid x_1 \rightsquigarrow_\delta x_0' \mid x_1 \rightsquigarrow_\gamma x_0'$. ∎

Lemma 28 Let $s \rightsquigarrow_\gamma^* t$ such that t is $\rightsquigarrow$-normal. If $s \rightsquigarrow_\delta u$ then $u \rightsquigarrow_\gamma^* t$.

Proof Observe that there must be a subterm $x_0 \mid x_i$ of s such that wlog. $s \rightsquigarrow_\delta u$ causes $x_0 \mid x_1 \rightsquigarrow_\delta x_0' \mid x_1$ and among the $s \rightsquigarrow_\gamma^* t$ there is the transition step $x_0 \mid x_1 \rightsquigarrow_\gamma x_1$. Otherwise, the δ-redex that $s \rightsquigarrow_\delta u$ unfolds would be preserved in t, contradicting its $\rightsquigarrow$-normality. Then $s \rightsquigarrow_\delta u \rightsquigarrow_\gamma^* t$ where $u \rightsquigarrow_\gamma^* t$ is “the same transitions” as $s \rightsquigarrow_\gamma^* t$ except the ones taking place within the subterm x_0 which might be disturbed by the changes introduced in $s \rightsquigarrow_\delta u$ but which are ultimately discarded by the transition $x_0' \mid x_1 \rightsquigarrow_\gamma x_1$. ∎

Theorem 29 Let s be an AHORS term and s' its $\rightsquigarrow_\delta$-normal form. Then any $\rightsquigarrow$-normal form t of s can be arrived at by a reduction sequence $s \rightsquigarrow_\delta^* s' \rightsquigarrow_\gamma^* t$.

Proof Suppose $s \rightsquigarrow^* t$ where t is $\rightsquigarrow$-normal. First observe that with repeated applications of Lemma 27, the reduction sequence $s \rightsquigarrow^* t$ can be transformed into $s \rightsquigarrow_\delta^* u \rightsquigarrow_\gamma^* t$ for some term u. By the uniqueness of normal forms, there is a reduction sequence $u \rightsquigarrow_\delta^* s'$. By repeated applications of Lemma 28, $u \rightsquigarrow_\delta^* s'$ and $u \rightsquigarrow_\gamma^* t$ together yield $s \rightsquigarrow_\delta^* u \rightsquigarrow_\delta^* s' \rightsquigarrow_\gamma^* t$ as desired. ∎

6 Language Size Bounds

6.1 Deriving an Upper Bound

In this section, we are only concerned with the γ-normalization of a term. For this, we consider terms AHORS in a slightly modified syntax in which each choice is annotated with a unique $i \in \omega$, that is, each choice is of the form $l_i \mid^i r_i$. We can then define a domination ordering, similar to that employed in the study of the modal μ-calculus.

Definition 30 Fix a choice annotated term s.

1. We define $\mathrm{Ix}(s) \coloneqq \{i \in \omega \mid l_i \mid^i r_i \text{ is a subterm of } s\}$, i.e. the set of choice-indices appearing in s
2. We write $i <_s j$ for $i, j \in \mathrm{Ix}(s)$ if $l_i \mid^i r_i$ is a subterm of l_j or r_j. If $i <_s j$ we say **j dominates i in s.**
3. We define a **domination side function** $\mathrm{ds} : \Pi i \in \mathrm{Ix}(s).\, \mathrm{Up}_{<_s}(i) \to \{l, r\}$ such that for $i <_s j$ we have $\mathrm{ds}(i, j) = l$ iff $l_i \mid^i r_i$ is a subterm of l_j

Lemma 31 Fix a choice annotated s. Then there exists an indexed family of linear orders $\prec_t$ for $t \in \mathrm{Up}_{\rightsquigarrow_\gamma^*}(s)$ such that $<_t \subseteq \prec_t$ and $\prec_t \subseteq \prec_u$ if $t \rightsquigarrow_\gamma^* u$.

Proof First of all, take $\prec_s$ to be some linearization of $<_s$. Such a linearization can always obtained as $<_s$ is a forest on $\mathrm{Ix}(s)$. For any $t \in \mathrm{Up}(s)$, simply take $\prec_t := \prec_s \cap \mathrm{Ix}(t)^2$. Notice that $<_t \subseteq \prec_t$ because $\rightsquigarrow_\gamma$-reductions do not disturb the domination order. ∎

We also employ an index-based notation for $\rightsquigarrow_\gamma$-reductions.

Definition 32 We write $s \rightsquigarrow_\gamma^{i,x} t$ for $i \in \mathrm{Ix}(s), x \in \{l, r\}$ for the $\rightsquigarrow_\gamma$-reduction which is justified by the γ-rule application $l_i \mid^i r_i \rightsquigarrow_\gamma x_i$. We sometimes write $s \rightsquigarrow_\gamma^i t$ if the $x \in \{l, r\}$ is not of import.

Lemma 33 Let $i \in \mathrm{Ix}(s)$ and $s \rightsquigarrow_\gamma^i t$ then $\mathrm{Up}_{\prec_s}(i) \subseteq \mathrm{Ix}(t)$.

Proof Observe that if $j \in \mathrm{Up}_{\prec_s}(i)$ then either $i < j$ or i and j are incomparable. In either case, the γ-step on i does not discard $l_j \mid^j r_j$, meaning $j \in \mathrm{Ix}(t)$. ∎

Corollary 34 Let $i <_s j$ and $s \rightsquigarrow_\gamma^j t \rightsquigarrow_\gamma^i u$ then there is t' such that $s \rightsquigarrow_\gamma^i t' \rightsquigarrow_\gamma^j u$.

We use this indexing mechanism to give a canonical representation of $\rightsquigarrow_\gamma$-reduction sequences originating at s. Note that the operation $s \Downarrow \sigma$ is well-defined by Lemma 33.

Definition 35 A **choice sequence of s** is a partial function $\sigma : \mathrm{Ix}(s) \rightharpoonup \{l, r\}$.

We recursively define a function $\Downarrow : \Pi s \in \mathrm{Tm}.\ (\mathrm{Ix}(s) \rightharpoonup \{l, r\}) \longrightarrow \mathrm{Tm}$ which is given by

$$s \Downarrow \sigma := \begin{cases} s & \sigma = \emptyset \\ t \Downarrow \sigma \restriction \mathrm{Up}_{\prec_s}(i) & i = \min_{\prec_s} \mathrm{Ix}(s) \text{ and } s \rightsquigarrow_\gamma^{i,\sigma(i)} t \end{cases}$$

These choice sequences contain enough information to describe every term reachable by some $\rightsquigarrow_\gamma$-sequence.

Lemma 36 Let $s \rightsquigarrow_\gamma^* t$ then there exists some $\sigma : \mathrm{Ix}(s) \rightharpoonup \{l, r\}$ such that $s \Downarrow \sigma = t$.

Proof Observe that $s \rightsquigarrow_\gamma^* t$ means there are $i_0, \ldots, i_n \in \mathrm{Ix}(s)$ and $x_0, \ldots, x_n \in \{l, r\}$ such that $s \rightsquigarrow_\gamma^{i_0,x_0} s_0 \rightsquigarrow_\gamma^{i_1,x_1} \ldots \rightsquigarrow_\gamma^{i_n,x_n} t$. Then take $\sigma := \{(i_l, x_l) \mid l \leq n\}$. To see that indeed $s \Downarrow \sigma = t$, observe that the reduction sequence $s \rightsquigarrow_\gamma^* t$ above can be rearranged into the one traced by $s \Downarrow \sigma$ using Corollary 34 while leaving the resulting term unchanged. ∎

We close our analysis of choice sequences with a result on the choice sequences of $\rightsquigarrow_\gamma$-normal terms.

Definition 37 An $i \in \mathrm{Ix}(s)$ is **discarded by σ** if there exists $i <_s j \in \mathrm{dom}(\sigma)$ such that $\sigma(j) \neq \mathrm{ds}(i, j)$.

Lemma 38 If $i \in \mathrm{Ix}(s)$ is discarded by σ then $s \Downarrow \sigma = s \Downarrow \sigma \cup \{(i, x)\}$ for $x \in \{l, r\}$.

Proof Let there be an $i <_s j \in \mathrm{dom}(\sigma)$ such that $\sigma(j) \neq \mathrm{ds}(i, j)$. That means the reduction traces of $s \Downarrow \sigma$ and $s \Downarrow \sigma \cup \{(i, x)\}$ eventually discard the subterm of $l_j \mid^j r_j$ containing $l_i \mid^i r_i$, making every previous reduction in said subterm irrelevant to the final term. ∎

Corollary 39 Let $s \rightsquigarrow_\gamma^* t$ and t be $\rightsquigarrow_\gamma$-normal. There exists a total σ with $s \Downarrow \sigma = t$.

Proof By Lemma 36 there exists a partial $\sigma' : \mathrm{Ix}(s) \dot{\rightarrow} \{l, r\}$ such that $s \Downarrow \sigma' = t$. Now observe that every $i \in \mathrm{Ix}(s) \setminus \mathrm{dom}(\sigma')$ needs to be discarded by σ'. After all, if some such i was not discarded, its redex would "survive" in $s \Downarrow \sigma'$, contradicting its $\rightsquigarrow_\gamma$-normality. Then define $\sigma := \sigma' \cup \{(i, l) \mid i \in \mathrm{Ix}(s) \setminus \mathrm{dom}(s)\}$ and observe that σ is total and by Lemma 38 $s \Downarrow \sigma = s \Downarrow \sigma' = t$ as desired. ∎

Based on this analysis, we can derive upper bounds on the language size of AHORS terms.

Theorem 40

1. For a term s we have $l(s) \leq 2^{d(s)}$
2. We obtain the following size- and height-based bounds

$$ls_n(N) \leq 2_{n+1}(N) \qquad lh_n(N) \leq 2_{n+2}(N)$$

3. We obtain the following complexity results

$$ls_n(N) = 2_{n+1}(O(n)) \qquad lh_n(N) = 2_{n+2}(N)$$

Proof

1. Recall that $L(s) = \{t \mid s \rightsquigarrow^* t, t \rightsquigarrow -\text{normal}\}$. By Theorem 29, any such $t \in L(s)$ can be arrived at via $s \rightsquigarrow_\delta^* s' \rightsquigarrow_\gamma^* t$ where s' is the $\rightsquigarrow_\delta$-normal form of s. That means $L(s) = L_\gamma(s')$. By Corollary 39, $L(s') \subseteq s' \Downarrow [\mathrm{Ix}(s') \rightarrow \{l, r\}]$. It is easily observed that $|\mathrm{Ix}(s') \rightarrow \{l, r\}| = 2^{|\mathrm{Ix}(s')|}$ and $|\mathrm{Ix}(s')| = d_\gamma(s')$, yielding the bound $L(s') \leq 2^{d_\gamma(s')}$. Now as $d_\gamma(s') \leq d(s') \leq d(s)$ we can conclude the desired bound $l(s) \leq 2^{d(s)}$.
2. This directly follows from 1. and the reduction length bounds in Theorem 13.
3. This directly follows from 2. ∎

6.2 Deriving a Lower Bound

The strategy to obtaining these lower bounds is very similar to those in Section 4.2: We find terms which have sufficient "blow up" and generate a lot of choice connectives, which in turn generate a big language. For this, we first need a result relating binary tree expressions with their resulting language size.

Lemma 41 Let $n \le l(s)$ then $n^{2^K} \le l(T_K(s))$.

Proof Proof per induction on K. The case for $K = 0$ is trivial. For the $K + 1$ case, observe that $L(T_{K+1}(s)) \supseteq \{b\,t\,t' \mid t, t' \in L(T_K(s))\}$, meaning

$$l(T_{k+1}(s)) \ge l(T_K(s))^2 \overset{\text{IH}}{\ge} (n^{2^K})^2 = n^{2^{K+1}}$$

∎

Theorem 42 There exist terms $L_n^s(N)$ and $L_n^h(N)$ with $s(L_n^s(N)) = h(L_n^h(N)) = O(N)$ and $\deg(L_n^s(N)) = \deg(L_n^h(N)) = n$. Furthermore,

$$l(L_n^s(N)) \ge 2_{n+1}(N) \qquad \text{and} \qquad l(L_n^h(N)) \ge 2_{n+2}(N)$$

Proof Fix terminals $a \ne b \in \mathcal{T}$. Then take $L_n^s(N) := S_n^N(a + b)$ and observe the fact that $L_n^s(N) \rightsquigarrow^* T_{2_{n-1}}(a + b)$ and thus $l(L_n^s(N)) \ge 2_{n+1}(N)$ by Theorem 42. Similarly, take $L_n^h(N) := H_n^N(a + b)$. ∎

Corollary 43

$$ls_n(N) \le 2_{n+2}(N) = 2_{n+2}(O(N)) \qquad lh_n(N) \le 2_{n+1}(N) = 2_{n+1}(O(N))$$

References

[1] Arnold Beckmann. "Exact bounds for lengths of reductions in typed λ-calculus". In: The Journal of Symbolic Logic 66.3 (Sept. 2001). Publisher: Cambridge University Press, pp. 1277–1285. ISSN: 0022-4812, 1943-5886. DOI: 10.2307/2695106. (Visited on 03/12/2021).