A Knuth-Bendix Like Ordering for Orienting Combinator Equations

Ahmed Bhayat and Giles Reger

University of Manchester, Manchester, UK

Abstract. Bláh-de-bláhs

1 Introduction

First-order theorem provers are often used as backends for interactive and higher-order theorem provers. The interactive provers translate their core logic into first-order logic which is then passed to first-order theorem provers. λ-expressions are often translated using combinators. It is well known that the set of combinators S, K and I are sufficient to translate any λ-expression. Thus, adding the three polyorphic axioms $\mathsf{S}\,x\,y\,z = x\,z\,(y\,z)$, $\mathsf{K}\,x\,y = x$ and $\mathsf{I}\,x = x$ to the translation of a higher-order problem results in the translation and the original problem being equi-satisfiable . wooly here

The reason why this has not, till now, been a viable method of automated theorem in higher-order logic, is that inferences between the axioms can be hugely explosive. Consider for example a superposition inference from the K axiom onto the right-hand side of the S axiom. The result is $\mathsf{S}\,\mathsf{K}\,y\,z = z$. There is little to restrict such inferences.

However, if an ordering could be found that orients the axioms left-to-right such inferences would become impossible. Superposition inferences only occur the maximal sides of equations and not at variables. A quick inspection of the left-hand side of the combinator axioms shows that if these are always the maximal sides, superposition between combinator axioms becomes impossible.

In this paper we investigate such an ordering. The ordering that we find lacks two monotonicity properties. It is not compatible with arguments. Bentkamp et al. [2] have already shown how the superposition calculus can be modified to circumvent this. Secondly, it is not compatible with contexts directly beneath combinators. In future work, we hope to show that this too can be circumvented.

2 Preliminaies

We work in a polymorphic first order logic. Let $\mathcal{V}_{\mathsf{ty}}$ be a set of type variables and Σ_{ty} be a set of type constructors with fixed arities. It is assumed that a binary type constructor $\to$ is present in Σ_{ty} which is written infix. The set of types is defined:

$$\textbf{Polymorphic Types} \quad \tau ::= \kappa(\,\overline{\tau_n}\,) \mid \alpha \mid \tau \to \tau \qquad \text{where } \alpha \in \mathcal{V}_{\mathsf{ty}} \text{ and } \kappa \in \Sigma_{\mathsf{ty}}$$

The notation $\overline{t_n}$ is used to denote a tuple or list of types or terms depending on the context. A type declaration is of the form $\Pi\,\overline{\alpha}\,.\sigma$ where σ is a type and all type variables in σ appear in $\overline{\alpha}$. Let Σ be a set of typed function symbols and $\mathcal{V}$ a set of variables with associated types. In what follows, $\mathsf{a}, \mathsf{b}, \mathsf{c}$ represent function symbols of arity 0 known as *constants*, $\mathsf{f}, \mathsf{g}, \mathsf{h}$ function symbols of non-zero arity and x, y, z variables. It is assumed that Σ contains the following function symbols, known as *basic combinators*:

How to introduce combinators nicely?

$$
\begin{aligned}
&\mathbf{S} : \Pi\alpha\tau\gamma.(\alpha \to \tau \to \gamma) \to (\alpha \to \tau) \to \alpha \to \gamma \\
&\mathbf{C} : \Pi\alpha\tau\gamma.(\alpha \to \tau \to \gamma) \to \tau \to \alpha \to \gamma \\
&\mathbf{B} : \Pi\alpha\tau\gamma.(\alpha \to \gamma) \to (\tau \to \alpha) \to \tau \to \gamma \\
&\mathbf{K} : \Pi\alpha\gamma.\alpha \to \gamma \to \alpha \\
&\mathbf{I} : \Pi\alpha.\alpha \to \alpha
\end{aligned}
$$

The set of terms over Σ and $\mathcal{V}$ is defined below. In what follows, type subscripts are generally omitted.

Terms $\quad \mathcal{T} ::= x \mid \mathsf{c} \mid t_{1_{\tau_1 \to \tau_2}} t_{2_{\tau_1}} \qquad$ where $\mathsf{c} \in \Sigma, x \in \mathcal{V}$ and $t_1, t_2 \in \mathcal{T}$

Following [1], terms of the form $t_1\, t_2$ are called applications. Non-application terms are called heads. A term can uniquely be decomposed into a head an n arguments. Let $t = \zeta\, \overline{t'_n}$. Then $head(t) = \zeta$ where ζ could be a variable or constant. The symbol $\mathcal{C}$ denotes an arbitrary combinator, the symbols $x, y, z \ldots$ are reserved for variables, $\mathsf{c}, \mathsf{d}, \mathsf{f} \ldots$ for non-combinator constants and ζ, ξ range over variables and non-combinator constants. For a term t, if $t \in \mathcal{V}$ or $t = \mathsf{f}$, then $pos(t) = \{\epsilon\}$. If $t = t_1 t_2$ then $pos(t) = \{\epsilon\} \cup_{1 \leq i \leq 2} \{i.p \mid p \in pos(t_i)\}$. Subterms at positions of the form $p.1$ are called *prefix* subterms and subterms at positions of the form $p.2$ are known as *first-order* subterms. A position p is *above* a position p' (denoted $p < p'$) if $\exists p''.p'' \neq \epsilon \;\wedge\; p' = p.p''$. Positions p and p' are *incomparable* (denoted $p \parallel p'$) if neither $p < p'$ nor $p' < p$. By $|t|$, the number of symbols occurring in t is denoted.

We define a subset of first-order subterms called *stable* subterms. Let $\mathsf{FNP}(t, p)$ be a partial function that takes a term t, a position p and returns the longest prefix p' of p such that $head(t|_{p'})$ is not a partially applied combinator if such a position exists. For a position $p \in pos(t)$, p is a stable position in t if $\mathsf{FNP}(t, p)$ is not defined or $head(t|_{\mathsf{FNP}(t,p)})$ is not a combinator. A *stable subterm* is a subterm occurring at a stable position and is denoted $t\langle\!\langle u \rangle\!\rangle|_p$. We call $t\langle\!\langle\rangle\!\rangle|_p$ a *stable context* and drop the position where it is not relevant. For example, the subterm a is not stable in $\mathsf{f}\,(\mathbf{S}\,\mathsf{a}\,\mathsf{b}\,\mathsf{c})$, $\mathbf{S}\,(\mathbf{S}\,\mathsf{a})\,\mathsf{b}\,\mathsf{c}$ and $\mathsf{a}\,\mathsf{c}$, but is in $\mathsf{g}\,\mathsf{a}\,\mathsf{b}$ and $\mathsf{f}\,(\mathbf{S}\,\mathsf{a})\,\mathsf{b}$. A subterm that is not stable is known as an *unstable* subterm.

This definition can be sharpened

The notation $t[u]$ denotes an arbitrary subterm u of t that may be unstable. The notation $t[u_1, \ldots, u_n]$, at times given as $t[\,\overline{u_n}\,]$ denotes that the the term t contains n *non-overlapping* subterms u_1 to u_n. By $u[]_n$, we refer to a context with n non-overlapping holes.

Terms have been defined over the set $\mathcal{V}$. However, our ordering works only on ground terms, so unless otherwise stated, from here on all terms are to be assumed ground. The ordering can be lifted to non-ground terms via a semantic lifting.

Each combinator is defined by its characteristic equation; $\mathsf{S}\,x\,y\,z = x\,z\,(y\,z)$, $\mathsf{C}\,x\,y\,z = x\,z\,y$, $\mathsf{B}\,x\,y\,z = x\,(y\,z)$, $\mathsf{K}\,x\,y = x$ and $\mathsf{I}\,x = x$. A term t weak-reduces to a term t' in one step (denoted $t \longrightarrow_w t'$) if $t = u[s]_p$ and there exists a combinator axiom $l = r$ and substitution σ such that $l\sigma = s$ and $t' = u[r\sigma]_p$. The term $l\sigma$ in t is called a *weak redex* or just redex. Weak-reduction is terminating and confluent as proved in [3]. By $(t)\!\downarrow^w$, we denote the term formed from t by contracting its leftmost redex.

The length of the longest weak-reduction from a term t is denoted $\|t\|$. This measure is one of the crucial features of the ordering. The next lemma states that a particular reduction strategy always produces the longest possible weak-reduction. It is in a sense equivalent to Barendregt's 'perpetual stratgey' in the λ-calculus.

Lemma 1. *Let F_∞ be a map from $\mathcal{T}$ to $\mathcal{T}$. Define F_∞ as follows:*

$$
\begin{array}{lcl}
F_\infty(t) & = & t \quad \textit{if } \|t\| = 0 \\
F_\infty(\mathsf{f}\ \overline{t_n}\,) & = & \mathsf{f}\,t_1 \ldots t_{i-1}\,F_\infty(t_i)\,t_{i+1} \ldots t_n \quad \textit{where } \|t_j\| = 0 \textit{ for } 1 \leq j < i \\
F_\infty(\mathsf{C}\,t_1\,t_2\,t_3) & = & (\mathsf{C}\,t_1\,t_2\,t_3)\!\downarrow^w \\
F_\infty(\mathsf{I}\,t') & = & t' \\
F_\infty(\mathsf{K}\,t_1\,t_2) & = & \begin{cases} t_1 \quad \textit{if } \|t_2\| = 0 \\ \mathsf{K}\,t_1\,F_\infty(t_2) \quad \textit{otherwise} \end{cases}
\end{array}
$$

The reduction strategy F_∞ is maximal.

Proof. Haven't worked out yet, but should be similar to Barendregt's proof of perpetual strategy or Raamsdonk et al.

3 Term Order

First Blanchette et al.'s [1] graceful higher-order KBO is presented as it is utilised within our ordering. The presentation here differs slightly from that in [1] because we do not allow ordinal weightings, all function symbols have finite arities and all terms are assumed ground.

3.1 Graceful Higher-Order KBO

Standard first-order KBO allows function symbols with weight 0. In a logic with partial application, this can be dangerous because it breaks the invariant that all terms have a minimum weight greater that some non zero-value ε. The graceful higher-order KBO is designed to solve this issue by introducing a penalty of ε for each missing argument of a function[1]. Thus, if f is a unary function of weight 0 the *term* f will have weight $\varepsilon \cdot arity(\mathsf{f}) = \varepsilon$. Following the authors of REF, these two notions of weight are represented by w and $\mathcal{W}$ respectively.

[1] The authors of [1] actually allow the penalty to range between 0 and ε, but the most useful case is when the penalty is ε, so we concentrate on that here

Let $\succ$ be a total well-founded ordering or *precedence* on Σ. Let w be a function from Σ to $\mathbb{N}$ and $\mathcal{W}$ a function from $\mathcal{T}$ to $\mathbb{N}$. Let $\varepsilon \in \mathbb{N}_{>0}$. For all constants c, $w(\mathsf{c}) \geq \varepsilon$. If, for some unary function ι, $w(\iota) = 0$, then $\iota \succ \mathsf{g}$ for all $\mathsf{g} \in \Sigma$. The weight of a term is defined recursively:

$$\mathcal{W}(\mathsf{f}) = w(\mathsf{f}) + arity(\mathsf{f}) \cdot \varepsilon \qquad \mathcal{W}(s\,t) = \mathcal{W}(s) + \mathcal{W}(t) - \varepsilon$$

The graceful higher-order Knuth-Bendix order $>_{\mathsf{hz}}$ is defined inductively as follows. Let $t = \zeta\ \bar{t}$ and $s = \xi\ \bar{s}$. Then $t >_{\mathsf{hz}} s$ if any of the following are satisfied:

Z1 $\mathcal{W}(t) > \mathcal{W}(s)$
Z2 $\mathcal{W}(t) = \mathcal{W}(s)$, $\bar{t} = t' >_{\mathsf{hz}} s$ and $\zeta = \iota$
Z3 $\mathcal{W}(t) = \mathcal{W}(s)$ and $\zeta \succ \xi$
Z4 $\mathcal{W}(t) = \mathcal{W}(s), \zeta = \xi$ and $\bar{t} \gg^{\mathsf{lex}}_{\mathsf{hz}} \bar{s}$

3.2 Combinator Orienting KBO

The combinator compatible KBO has the property that all combinator axioms are oriented left to right. This is achieved by first comparing terms by the length of the longest weak-reduction from the term and then using $>_{\mathsf{hz}}$. In Section 4 some useful properties of the resulting ordering are proved.

The combinator orienting Knuth-Bendix order $>_{\mathsf{ski}}$ is defined as follows. For terms t and s, $t >_{\mathsf{ski}} s$ if:

R1 $\|t\| > \|s\|$ or,
R2 $\|t\| = \|s\|$ and $t >_{\mathsf{hz}} s$

Lemma 2. *For all combinator axioms or instances of them $l \approx r$, it is the case that $l >_{\mathsf{ski}} r$.*

Proof. Since for all axioms of instances of the axioms $l \approx r$, we have $\|l\| > \|r\|$, the theorem follows by an application of R1.

4 Properties

Various properties of the order $>_{\mathsf{ski}}$ are proved here. Note that the order is not a reduction ordering as it lacks compatibility with contexts.

Theorem 1 (Irreflexivity). *For all terms s, it is not the case that $s >_{\mathsf{ski}} s$.*

Proof. It is obvious that $\|s\| = \|s\|$. Therefore $s >_{\mathsf{ski}} s$ can only be derived by rule R2. However, this is precluded by the irreflexivity of $>_{\mathsf{hz}}$.

Theorem 2 (Transitivity). *For terms s, t and u, if $s >_{\mathsf{ski}} t$ and $t >_{\mathsf{ski}} u$ then $s >_{\mathsf{ski}} u$*

Proof. If $\|s\| > \|t\|$ or $\|t\| > \|u\|$ then $\|s\| > \|u\|$ and $s >_{\mathsf{ski}} u$ follows by an application of rule R1. Therefore, suppose that $\|s\| = \|t\| = \|u\|$. Then it must be the case that $s >_{\mathsf{hz}} t$ and $t >_{\mathsf{hz}} u$. It follows from the transitivity of $>_{\mathsf{hz}}$ that $s >_{\mathsf{hz}} u$ and thus $s >_{\mathsf{ski}} u$.

Theorem 3 (Ground Totality). *Let s and t be ground terms that are not syntactically equal. Then either $s >_{\mathsf{ski}} t$ or $t >_{\mathsf{ski}} s$.*

Proof. If $\|s\| \neq \|t\|$ then by R1 either $s >_{\mathsf{ski}} t$ or $t >_{\mathsf{ski}} s$. Otherwise, s and t are compared using $>_{\mathsf{hz}}$ and either $t >_{\mathsf{hz}} s$ or $s >_{\mathsf{hz}} t$ holds by the ground totality of $>_{\mathsf{hz}}$.

Theorem 4 (Subterm Property). *If t is a proper subterm of s then $s >_{\mathsf{ski}} t$.*

Proof. For all subterms u of s, it holds that $\|s\| \geq \|u\|$ because any weak-reduction in u is also a weak-reduction in s. If $\|s\| > \|u\|$, the theorem follows by an application of R1. Otherwise s and u are compared using $>_{\mathsf{hz}}$ and $s >_{\mathsf{hz}} u$ holds by the subterm property of $>_{\mathsf{hz}}$.

Lemma 3. $\|\zeta\,\overline{t_n}\,\| = \sum_{i=1}^{i=n} \|t_i\|$ *if ζ is not a fully applied combinator*

Lemma 4. *Let $\overline{t_n}$ be terms such that for each t_i, $head(t_i) \notin \{\mathsf{I}, \mathsf{K}, \mathsf{B}, \mathsf{C}, \mathsf{S}\}$. Let $\overline{t'_n}$ be terms with the same property. Moreover, let $\|t_i\| \geq \|t'_i\|$ for $1 \leq i \leq n$. Let $s = u[\,\overline{t_n}\,]$ and $s' = u[\,\overline{t'_n}\,]$ where each t_i and t'_i is at position p_i in s and s'. If the F_∞ redex in s is within t_i for some i, then the F_∞ redex in s' is within t'_i unless t'_i is in normal form.*

Proof. Proof is by induction on $|s| + |s'|$. If u is the empty context, then $n = 1$, $s = t_1$ and $s' = t'_1$. If $\|t'_1\| > 0$, the F_∞ redex of s' must be a subterm of t'_1 as $head(t'_1)$ is not a combinator.

If u is not the empty context, then $s = \mathsf{f}\ \overline{s_n}$ and $s' = \mathsf{f}\ \overline{s'_n}$. From the definition of F_∞, we have $F_\infty(s) = \mathsf{f}\, s_1 \ldots s_{i-1}\, F_\infty(s_i)\, s_{i+1} \ldots s_n$ where $\|s_j\| = 0$ for $1 \leq j < i$. Using $\|s_j\| = 0$ for $1 \leq j < i$ and $\|t_i\| \geq \|t'_i\|$ for $1 \leq i \leq n$, we have $F_\infty(s') = \mathsf{f}\, s'_1 \ldots s'_{i-1}\, F_\infty(s'_i)\, s'_{i+1} \ldots s'_n$. Let the F_∞ redex of s_i occur inside t_i. Then t'_i is a subterm of s'_i and by the induction hypothesis, the F_∞ redex of s'_i occurs inside t'_i. The lemma follows immediately.

Lemma 5. *Let $\overline{t_n}$ be terms such that for $1 \leq i \leq n$, $head(t_i) \notin \{\mathsf{I}, \mathsf{K}, \mathsf{B}, \mathsf{C}, \mathsf{S}\}$. Then for all contexts $u[]_n$, if $u[t_1, \ldots, t_n]_n \longrightarrow_w u'$ then either:*

1. *$\exists i. u' = u[t_1, \ldots, \hat{t}_i, \ldots, t_n]$ where $t_i \longrightarrow_w \hat{t}_i$ or*
2. *$u' = \hat{u}\{x_1 \to t_1, \ldots, x_n \to t_n\}$ where $u[x_1, \ldots, x_n] \longrightarrow_w \hat{u}$*

Proof. Let $s = u[\,\overline{t_n}\,]$ and let $p_1, \ldots p_n$ be the positions of $\overline{t_n}$ in s. Since s is reducible, there must exist a p such that $s|_p$ is a redex. If $p = p_i.p'$ where $p' \neq \epsilon$, then $u[t_1, \ldots, \hat{t}_i, \ldots, t_n]|_p = t_i|'_p \longrightarrow_w$

For every p_i, either $p < p_i$, $p > p_i$ or the two are incomparable. We cannot have $p = p_i$ because $head(t_i) \notin \{\mathsf{I}, \mathsf{K}, \mathsf{B}, \mathsf{C}, \mathsf{S}\}$ whilst $head(s) \in \{\mathsf{I}, \mathsf{K}, \mathsf{B}, \mathsf{C}, \mathsf{S}\}$. If $p > p_i$ for some i, then for all $j \neq i$, p_i and p_j are incomparable as the $\overline{p_n}$ are independent. Assume that $p > p_i$ for some i. Then there exists a p' such that $p_i.p' = p$, $s|_{p_i} = t_i$ and $t_i|_{p'} = s|_p$.

notation

not sure how to finish this off

Lemma 6. *Let $\overline{t_n}$ be terms such that for each t_i, $head(t_i) \notin \{\mathsf{I}, \mathsf{K}, \mathsf{B}, \mathsf{C}, \mathsf{S}\}$. Let $\overline{t'_n}$ be terms with the same property. Then:*

1. *If $\|t_i\| = \|t'_i\|$ for all i in $\{1, \ldots, n\}$, then $\|u[\,\overline{t_n}\,]_n\| = \|u[\,\overline{t'_n}\,]_n\|$ for all n holed contexts u.*

2. *If $\|t_i\| > \|t'_i\|$ for at least one i in $\{1, \ldots, n\}$, then $\|u[\,\overline{t_n}\,]_n\| > \|u[\,\overline{t'_n}\,]_n\|$ for all n holed contexts u.*

Proof. Proof is by induction on $\|s\| + \|s'\|$. Let $p_1, \ldots, p_n$ be the positions of the 't_i's and 't'_i's in u and let $s = u[\,\overline{t_n}\,]$ and $s' = u[\,\overline{t'_n}\,]$. We prove part (1) first:

Assume that $\|u[\,\overline{t_n}\,]\| = 0$. Then $\|t_i\| = 0$ for $1 \leq i \leq n$. Now assume that $\|u[\,\overline{t'_n}\,]\| \neq 0$. Then there must exist some position p such that $s'|_p$ is a redex. We have that $p \neq p_i$ for all p_i as $head(t'_i) \notin \{\mathsf{I}, \mathsf{K}, \mathsf{B}, \mathsf{C}, \mathsf{S}\}$. Assume $p > p_i$ for some p_i. But then, $\|t'_i\| > 0$ which contradicts the fact that $\|t_i\| = \|t'_i\|$ for all i. Therefore, for all p_i either $p < p_i$ or $p \parallel p_i$. But then, if $s'|_p$ is a redex, so must $s|_p$ be, contradicting the fact that $\|u[\,\overline{t_n}\,]\| = 0$. Thus, we conclude that $\|u[\,\overline{t'_n}\,]\| = 0$.

Problem. For $p < p_i \vee p \parallel p_i$, we do **not** have $head(s|_p) = head(s'|_p)$.

Assume that $\|u[\,\overline{t_n}\,]\| > 0$. Let $u' = F_\infty(s)$. By Lemma 5 either $u' = u[t_1, \ldots, \hat{t_i}, \ldots, t_n]$ where $t_i \longrightarrow_w \hat{t_i}$ for $0 \leq i \leq n$ or $u' = \hat{u}\{\,\overline{x_n} \rightarrow \overline{t_n}\,\}$ where $u[\,\overline{x_n}\,] \longrightarrow_w \hat{u}$. In the first case, by Lemma 4 and $\|t_i\| = \|t'_i\|$ we have $F_\infty(s') = u'' = u[t'_1, \ldots, \hat{t'_i}, \ldots, t'_n]$ where $t'_i \longrightarrow_w \hat{t'_i}$. By the induction hypothesis $\|u'\| = \|u''\|$ and thus $\|s\| = \|s'\|$. In the second case, $F_\infty(s') = u'' = \hat{u}\{\,\overline{x_n} \rightarrow \overline{t'_n}\,\}$ where $u[\,\overline{x_n}\,] \longrightarrow_w \hat{u}$. Again, the induction hypothesis can be used to show $\|u\| = \|u''\|$ and the theorem follows.

We now prove part (2); $\|u[\,\overline{t_n}\,]\|$ must be greater than 0. Again, let $u' = F_\infty(s)$ and $u'' = F_\infty(s')$. If $u' = u[t_1, \ldots, \hat{t_i}, \ldots, t_n]$ and $\|t'_i\| \neq 0$, then by Lemma 4 $u'' = u[t'_1, \ldots, \hat{t'_i}, \ldots, t'_n]$ where $t'_i \longrightarrow_w \hat{t'_i}$. The lemma follows by the induction hypothesis. If $\|t'_i\| = 0$, consider terms u' and s'. If $\|\hat{t_i}\| > 0$ or $\|t_j\| > \|t'_j\|$ for some $j \neq i$, then the induction hypothesis can be used to show $\|u'\| > \|s'\|$ and therefore $\|s\| = \|u'\| + 1 > \|s'\|$. Otherwise, $\|t_j\| = \|t'_j\|$ for all $j \neq i$ and $\|\hat{t_i}\| = 0 = \|t'_i\|$. Part 1 of this lemma can be used to show that $\|u'\| = \|s\|$ and thus $\|s\| = \|u'\| + 1 > \|s'\|$. If $u' = \hat{u}\{\,\overline{x_n} \rightarrow \overline{t_n}\,\}$, then $u'' = \hat{u}\{\,\overline{x_n} \rightarrow \overline{t'_n}\,\}$ and the lemma follows by the induction hypothesis.

Lemma 7 (Compatibility with Contexts). *For terms s and t, such that $head(s), head(t) \notin \{\mathsf{I}, \mathsf{K}, \mathsf{B}, \mathsf{C}, \mathsf{S}\}$, and $s >_{\mathsf{ski}} t$, then $u[s] > u[t]$.*

Proof. By Lemma 6, we have that if $\|s\| > \|t\|$, then $\|u[s]\| > \|u[t]\|$. Thus, if $s >_{\mathsf{ski}} t$ was derived by R1, $u[s] >_{\mathsf{ski}} u[t]$ follows by R1. Otherwise, $s >_{\mathsf{ski}} t$ is derived by R2 and $\|s\| = \|t\|$. By Lemma 6, $\|u[s]\| = \|u[t]\|$ follows. Thus, $u[s]$ is compared with $u[t]$ by R2 and $u[s] > u[t]$ by the compatibility with contexts of $>_{\mathsf{hz}}$.

Lemma 8. $\|s\| > \|t\| \implies \|u\langle\!\langle s\rangle\!\rangle\| > \|u\langle\!\langle t\rangle\!\rangle\|$ *and* $\|s\| = \|t\| \implies \|u\langle\!\langle s\rangle\!\rangle\| = \|u\langle\!\langle t\rangle\!\rangle\|$

Proof. Proceed by induction on the size of the context u. If u is the empty context, both parts of the theorem hold trivially.

The inductive case is proved for part (1) of the theorem first. If u is not the empty context, $u\langle\!\langle s\rangle\!\rangle$ is of the form $u'\langle\!\langle \zeta\, t_1 \ldots t_{i-1}, s, t_{i+1} \ldots t_n\rangle\!\rangle$. By the definition of a stable subterm ζ cannot be a fully applied combinator and thus by Lemma 3 we have $\|\zeta\, t_1 \ldots t_{i-1}, s, t_{i+1} \ldots t_n\| = \sum_{i=1}^{i=n} \|t_i\| + \|s\| > \sum_{i=1}^{i=n} \|t_i\| + \|t\| = \|\zeta\, t_1 \ldots t_{i-1}, t, t_{i+1} \ldots t_N\|$. If ζ is not a combinator, then $\|u'\langle\!\langle \zeta\, t_1 \ldots t_{i-1}, s, t_{i+1} \ldots t_n\rangle\!\rangle\| = \|u'\langle\!\langle \zeta\, t_1 \ldots t_{i-1}, t, t_{i+1} \ldots t_n\rangle\!\rangle\|$ follows from Lemma 6. Otherwise, as u' is smaller stable context than u, the induction hypothesis can be used to conclude that $\|u'\langle\!\langle \zeta\, t_1 \ldots t_{i-1}, s,$

$t_{i+1} \ldots t_n\rangle\!\rangle\| = \|u'\langle\!\langle \zeta\, t_1 \ldots t_{i-1}, t, t_{i+1} \ldots t_n\rangle\!\rangle\|$ and thus that $\|u\langle\!\langle s\rangle\!\rangle\| = \|u\langle\!\langle t\rangle\!\rangle\|$. The proof of the inductive case for part (2) is almost identical.

Theorem 5 (Compatibility with Stable Contexts). *For all stable contexts $u\langle\!\langle\rangle\!\rangle$, if $s >_{\mathsf{ski}} t$ then $u\langle\!\langle s\rangle\!\rangle >_{\mathsf{ski}} u\langle\!\langle t\rangle\!\rangle$*

Proof. If $\|s\| > \|t\|$ then by Lemma 8, $\|u\langle\!\langle s\rangle\!\rangle\| > \|u\langle\!\langle t\rangle\!\rangle\|$ holds and then by an application of R1 we have $u\langle\!\langle s\rangle\!\rangle >_{\mathsf{ski}} u\langle\!\langle t\rangle\!\rangle$. Otherwise, if $\|s\| = \|t\|$, then by Lemma 8 we have that $\|u\langle\!\langle s\rangle\!\rangle\| = \|u\langle\!\langle t\rangle\!\rangle\|$. Thus $u\langle\!\langle s\rangle\!\rangle$ and $u\langle\!\langle t\rangle\!\rangle$ are compared using $>_{\mathsf{hz}}$. By the compatibility with functions of $>_{\mathsf{hz}}$, $u\langle\!\langle s\rangle\!\rangle >_{\mathsf{hz}} u\langle\!\langle t\rangle\!\rangle$ holds and then by an application of R2 $u\langle\!\langle s\rangle\!\rangle >_{\mathsf{ski}} u\langle\!\langle t\rangle\!\rangle$ is true.

Theorem 6 (Compatibility with Arguments). *If $s >_{\mathsf{ski}} t$ and $head(s)$ and $head(t)$ are not combinators then $s\, u >_{\mathsf{ski}} t\, u$.*

Proof. Let $s = \zeta\ \overline{s'_n}$ and $t = \xi\ \overline{t'_m}$. Using Lemma 3 we have that $\|s\| = \sum_{i=1}^{i=n} \|s'_i\|$ and $\|t\| = \sum_{i=1}^{i=m} \|t'_i\|$. Thus, if $s >_{\mathsf{ski}} t$ has been derived by R1 then $\sum_{i=1}^{i=n} \|s'_i\| > \sum_{i=1}^{i=m} \|t'_i\|$. Using this inequality we have $\|s\, u\| = \sum_{i=1}^{i=n} \|s'_i\| + \|u\| > \sum_{i=1}^{i=m} \|t'_i\| + \|u\| = \|t\, u\|$. Therefore, $s\, u >_{\mathsf{ski}} t\, u$.

On the other hand, if $s >_{\mathsf{ski}} t$ has been derived by R2, then $\|s\, u\| = \|t\, u\|$ and $s\, u >_{\mathsf{ski}} t\, u$ follows by the compatibility with arguments of $>_{\mathsf{hz}}$.

Theorem 7 (Well-foundedness). *There exists no infinite descending chain of comparisons $s_1 >_{\mathsf{ski}} s_2 >_{\mathsf{ski}} s_3 \cdots$.*

Proof. Assume that such a chain exists. For each $s_i >_{\mathsf{ski}} s_{i+1}$ derived by R1, we have that $\|s_i\| > \|s_{i+1}\|$. For each $s_i >_{\mathsf{ski}} s_{i+1}$ derived by R2, we have that $\|s_i\| = \|s_{i+1}\|$. Therefore the number of times $s_i >_{\mathsf{ski}} s_{i+1}$ by R1 in an any infinite chain must be finite and there must exist some m such that for all $n > m$, $s_n >_{\mathsf{ski}} s_{n+1}$ by R2. Therefore, there exists an infinite sequence of $>_{\mathsf{hz}}$ comparisons $s_m >_{\mathsf{hz}} s_{m+1} >_{\mathsf{hz}} s_{m+2} \cdots$. This contradicts the well-foundedness of $>_{\mathsf{hz}}$.

References

1. Heiko Becker, Jasmin Christian Blanchette, Uwe Waldmann, and Daniel Wand. A transfinite knuth–bendix order for lambda-free higher-order terms. In Leonardo de Moura, editor, *Automated Deduction – CADE 26*, pages 432–453, Cham, 2017. Springer International Publishing.
2. Alexander Bentkamp, Jasmin Christian Blanchette, Simon Cruanes, and Uwe Waldmann. Superposition for lambda-free higher-order logic. In Didier Galmiche, Stephan Schulz, and Roberto Sebastiani, editors, *Automated Reasoning*, pages 28–46, Cham, 2018. Springer International Publishing.
3. J. Roger Hindley and Jonathan P. Seldin. *Lambda-Calculus and Combinators: An Introduction*. Cambridge University Press, New York, NY, USA, 2 edition, 2008.