A Knuth-Bendix-Like Ordering for Orienting Combinator Equations

Ahmed Bhayat and Giles Reger

University of Manchester, Manchester, UK

Abstract. We extend the graceful higher-order basic Knuth-Bendix order (KBO) of Blanchette et al. to an ordering that orients combinator equations left-to-right. The resultant ordering is highly suited to parameterising the first-order superposition calculus when dealing with the theory of higher-order logic, as it prevents inferences between the combinator axioms. We prove a number of desirable properties about the ordering including it having the subterm property for ground terms, being transitive and being well-founded. The ordering fails to be a reduction ordering as it lacks compatibility with certain contexts. We provide an intuition of why this need not be an obstacle when using it to parameterise superposition.

1 Introduction

There exists a wide range of methods for automated theorem proving in higher-order logic. Some provers such as AgsyHOL [?], Satallax [?] and Leo-II [?] implement dedicated higher-order proof calculi. A common approach, followed by the Leo-III prover [?], is to use a co-operative architecture with a dedicated higher-order prover working in conjunction with a first-order prover. It has long been part of theorem proving folklore that sound and complete translations from higher-order to first-order logic exist. Kerber [?] proves this result for a higher-order logic that does not assume comprehension axioms (otherwise known as applicative first-order logic). Thus, translating higher-order problems to first-order logic and running first-order provers on the translations is another method of automated higher-order theorem proving. Variations of this method are widely utilised by interactive theorem provers and their hammers such as Sledgehammer [?] and the CoqHammer [?].

Almost all translations to first-order logic translate λ-express-ions using combinators. It is well known that the set of combinators S, K and I are sufficient to translate any λ-expression. For purposes of completeness, these combinators must be axiomatised: $\mathsf{S}\langle\tau_1,\tau_2,\tau_3\rangle\, x\, y\, z = x\, z\, (y\, z)$, $\mathsf{K}\langle\tau_1,\tau_2\rangle\, x\, y = x$ and $\mathsf{I}\langle\tau\rangle\, x = x$. If translating to a monomorphic logic a finite set of axioms cannot achieve completeness.

However, till now, translation based methods have proven disappointing and only achieved decent results with interactive theorem provers when the problems are first-order or nearly first-order [?]. One of the major reasons why utilising translations has not proven a viable method of automated theorem proving in higher-order logic, is that inferences between the axioms are hugely explosive. A common first-order proof calculus is superposition [?]. Consider 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.

Superposition is parameterised by a simplification ordering and, in general, inferences are only carried out on the larger side of maximal literals with respect to this ordering. Thus, if a simplification ordering exists that orients the axioms left-to-right, inferences amongst the axioms would be impossible. A quick inspection of the left-hand side of the combinator axioms shows that none contain non-variable subterms that can be unified with the left-hand side of other axioms. Consequently, if the left-hand sides are always larger than the right by some simplification ordering, there can be no inferences amongst the axioms.

Currently, no such simplification ordering is known to exist and the authors suspect that no such ordering can exist. Whilst there is a large body of work on higher-order orderings, all either lack some property required for them to be simplifcation orderings or are unsuitable for orienting the combinator axioms. Jouannaud and Rubio introduced a higher-order version of the recursive path order called HORPO [**?**]. HORPO is compatible with β-reduction which suggests that without much difficulty it could be modified to be compatible with weak reduction. However, the ordering does not enjoy the subterm property, nor is it transitive. Likewise is the case for orderings based on HORPO such as the computability path ordering [**?**] and the iterative HOIPO of Kop and van Raamsdonk [**?**]. More recently, Blanchette et al. have developed a pair of orderings for λ-free higher-order terms [**?**][**?**]. These orderings lack a specific monotonicity property, but this does not prevent their use in superposition [**?**]. However, neither ordering orients combinator axioms directly.

In this paper we investigate an extension of the graceful higher-order basic KBO $>_{\mathsf{hb}}$ introduced by Blanchette et al. [**?**] that orients combinator equations. This allows our new ordering, $>_{\mathsf{ski}}$, to be used directly in a first-order superposition prover reasoning over a combinator-based translation of a higher-order problem, with the key advantage that combinator axioms will never be combined. We note that the $>_{\mathsf{ski}}$ ordering lacks two monotonicity properties. Firstly, it is not compatible with arguments. Bentkamp et al. [**?**] have already shown how the superposition calculus can be modified to remain complete despite this lack of monotonicity. Secondly, it is not compatible with contexts directly beneath combinators. In future work, we hope to show that this too can be circumvented.

In Section 2, we provide the necessary preliminaries and then move on to the main contributions of this paper which are:

- Two approaches to extending the $>_{\mathsf{hb}}$ ordering by first comparing terms by the length of the longest weak reduction from them. The first approach is simple to define but is conservative when comparing non-ground terms. The second approach relaxes some conditions allowing the comparison of more non-ground terms. (Section 3)
- A set of proofs that the introduced $>_{\mathsf{ski}}$ ordering enjoys the necessary properties required for its use within the superposition calculus (Section 4) and a set of examples demonstrating how the ordering applies to certain terms (Section 5).
- An extension of the previous idea to λ-*terms* (those containing λ-abstractions) with β-reduction replacing weak reduction (Section 6).

In our concluding remarks (Section 7), we discuss the possible advantages of parameterising the first-order superposition calculus with the $>_{\mathsf{ski}}$ ordering over the only other existing complete superposition calculus for higher-order logic that we are aware of. We also discuss methods of strengthening some of the results of this paper.

2 Preliminaries

Syntax of types and terms: we work in a polymorphic applicative 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 $\rightarrow$ is present in Σ_{ty} which is written infix. The set of types is defined:

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

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. It is assumed that Σ contains the following function symbols, known as *basic combinators*:

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

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

$$\textbf{Terms} \quad \mathcal{T} ::= x \mid \mathsf{f}\langle\,\overline{\tau_n}\,\rangle \mid t_{1\tau_1 \rightarrow \tau_2} t_{2\tau_1}$$
$$\text{where } x \in \mathcal{V},\, t_1, t_2 \in \mathcal{T},\, \mathsf{f} \in \Sigma,\, \mathsf{f} : \Pi\,\overline{\alpha_n}\,.\sigma \text{ and } \overline{\tau_n} \text{ are types}$$

The type of the term $\mathsf{f}\langle\,\overline{\tau_n}\,\rangle$ is $\sigma\{\,\overline{\alpha_n} \rightarrow \overline{\tau_n}\,\}$. Following [**?**], 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 and n arguments. Let $t = \zeta\;\overline{t'_n}$. Then $head(t) = \zeta$ where ζ could be a variable or constant applied to possibly zero type arguments. The symbol $\tilde{\mathcal{C}}$ denotes a member of $\{\mathsf{S}, \mathsf{C}, \mathsf{B}, \mathsf{K}, \mathsf{I}\}$, whilst $\mathcal{C}$ denotes a member of $\{\mathsf{S}, \mathsf{C}, \mathsf{B}\}$. These symbols are only used when the combinator is assumed to have a full complement of arguments. Thus, in $\mathcal{C}\;\overline{t_n}$, $n \geq 3$ is assumed. 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 arbitrary symbols and, by an abuse of notation, at times even terms. A term is *ground* if it contains no variables and *term ground* if it contains no term variables.

Positions over terms: for a term t, if $t \in \mathcal{V}$ or $t = \mathsf{f}\langle\,\overline{\tau}\,\rangle$, then $pos(t) = \{\epsilon\}$ (type arguments have no position). If $t = t_1 t_2$ then $pos(t) = \{\epsilon\} \cup \{i.p \mid 1 \leq i \leq 2,\; p \in$

$pos(t_i)\}$. 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 *strictly 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 $p' < p$, nor $p = p'$. By $|t|$, the number of symbols occurring in t is denoted. By $vars_{\#}(t)$ the multiset of variables in t is denoted. The expression $A \subseteq B$ means that either A is a subset of B or A is a submultiset of B depending on whether A and B are sets or multisets.

Stable subterms: 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 and 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 $\mathsf{f}\,(\mathbf{S}\,\mathsf{a}\,\mathsf{b}\,\mathsf{c})$, $\mathbf{S}\,(\mathbf{S}\,\mathsf{a})\,\mathsf{b}\,\mathsf{c}$ (in both cases, $head(t|_{\mathsf{FNP}(t,p)}) = \mathbf{S}$) and $\mathsf{a}\,\mathsf{c}$ (a is not a first-order subterm), 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.

The notation $t[u]|_p$ denotes an arbitrary subterm u of t that occurs at position p and 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. Whilst this resembles the notation for a term at position n, ambiguity is avoided by never using n to denote a position or p to denote a natural number.

Weak reduction: each combinator is defined by its characteristic equation; $\mathbf{S}\,x\,y\,z = x\,z\,(y\,z)$, $\mathbf{C}\,x\,y\,z = x\,z\,y$, $\mathbf{B}\,x\,y\,z = x\,(y\,z)$, $\mathbf{K}\,x\,y = x$ and $\mathbf{I}\,x = x$. 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. By $\longrightarrow_w^*$, the reflexive transitive closure of $\longrightarrow_w$ is denoted. If term t weak-reduces to term t' in n steps, we write $t \longrightarrow_w^n t'$. Further, if there exists a weak-reduction path from a term t of length n, we say that $t \in n_w$. Weak-reduction is terminating and confluent as proved in [**?**]. 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 investigated in this paper.

2.1 A Maximal Weak-reduction Strategy

To show that the measure $\|\,\|$ is computable we provide a maximal weak-reduction strategy and prove its maximality. The strategy is used in a number of proofs later in the paper. It is in a sense equivalent to Barendregt's 'perpetual strategy' in the λ-calculus [**?**]. Our proof of its maximality follows the style of van Raamsdonk et al. [**?**] in their proof of the maximality of a particular β-reduction strategy. We begin by proving the fundamental lemma of maximality for combinatory terms.

Lemma 1 (Fundamental Lemma of Maximality). $\|\tilde{\mathcal{C}}\ \overline{t_n}\,\| = \|(\tilde{\mathcal{C}}\ \overline{t_n}\,)\downarrow^w\ \| + 1 + isK(\tilde{\mathcal{C}}) \times \|t_1\|$ *where* $isK(\tilde{\mathcal{C}}) = 1$ *if* $\tilde{\mathcal{C}} = \mathbf{K}$ *and is* 0 *otherwise. The lemma holds for* $n \geq 3$ *if* $\tilde{\mathcal{C}} \in \{\mathbf{S}, \mathbf{C}, \mathbf{B}\}$*,* $n \geq 2$ *if* $\tilde{\mathcal{C}} = \mathbf{K}$ *and* $n \geq 1$ *otherwise.*

Proof. Assume that $\tilde{\mathcal{C}} = \mathsf{K}$. Then any maximal reduction from $\mathsf{K}\ \overline{t_n}$ is of the form:

$$\begin{array}{rlr} \mathsf{K}\, t_0\, t_1 \ldots t_n & \longrightarrow_w^m & \mathsf{K}\, t_0'\, t_1' \ldots t_n' \\ & \longrightarrow_w & t_0'\, t_2' \ldots t_n' \\ & \longrightarrow_w^{m'} & s \end{array}$$

where $\|s\| = 0$, $t_0 \longrightarrow_w^{m_0} t_0' \ldots t_n \longrightarrow_w^{m_n} t_n'$, $\|t_1\| = m_1$ and $m = m_0 + \ldots + m_n$. Thus, $\|\mathsf{K}\ \overline{t_n}\ \| = \sum_{i=0}^n m_i + 1 + m'$. There is another method of reducing $\mathsf{K}\ \overline{t_n}$ to s:

$$\begin{array}{rlr} \mathsf{K}\, t_0\, t_1 \ldots t_n & \longrightarrow_w^{m_1} & \mathsf{K}\, t_0\, t_1' \ldots t_n \\ & \longrightarrow_w & t_0\, t_2 \ldots t_n \\ & \longrightarrow_w^{m-m_1} & t_0'\, t_2' \ldots t_n' \\ & \longrightarrow_w^{m'} & s \end{array}$$

As the length of this reduction is the same as the previous reduction, it must be a maximal reduction as well. Therefore we have that:

$$\begin{array}{rlr} \|\mathsf{K}\, t_0\, t_1 \ldots t_n\| & = & m + m' + 1 \\ & = & (m - m_1 + m') + m_1 + 1 \\ & = & \|t_0\, t_2 \ldots t_n\| + \|t_1\| + 1 \end{array}$$

Conversely, assume that $\tilde{\mathcal{C}}$ is not K. We prove that the formula holds if $\tilde{\mathcal{C}} = \mathsf{S}$. The other cases are similar. If $\tilde{\mathcal{C}} = \mathsf{S}$, any maximal reduction from $\mathsf{S}\ \overline{t_n}$ must be of the form:

$$\begin{array}{rlr} \mathsf{S}\, t_0 \ldots t_n & \longrightarrow_w^m & \mathsf{S}\, t_0' \ldots t_n' \\ & \longrightarrow_w & t_0'\, t_2'\, (t_1'\, t_2')\, t_3' \ldots t_n' \\ & \longrightarrow_w^{m'} & s \end{array}$$

where $\|s\| = 0$, $t_0 \longrightarrow_w^{m_0} t_0' \ldots t_n \longrightarrow_w^{m_n} t_n'$ and $m = m_0 + \ldots + m_n$. There is another method of reducing $\mathsf{S}\ \overline{t_n}$ to s:

$$\begin{array}{rlr} \mathsf{S}\, t_0 \ldots t_n & \longrightarrow_w & t_0\, t_2\, (t_1\, t_2)\, t_3 \ldots t_n \\ & \longrightarrow_w^{m+m_2} & t_0'\, t_2'\, (t_1'\, t_2')\, t_3' \ldots t_n \\ & \longrightarrow_w^{m'} & s \end{array}$$

Thus, we have that $\|\mathsf{S}\ \overline{t_n}\ \| = m + m' + 1 \leq m + m_2 + m' + 1 = \|(\mathsf{S}\ \overline{t_n}\ ){\downarrow^w}\ \| + 1$. Since $m + m' + 1$ is the length of the maximal reduction, equality must hold.

Lemma 2. *Let F_∞ be a map from $\mathcal{T}$ to $\mathcal{T}$. Define F_∞ as follows:*

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

The reduction strategy F_∞ is maximal.

Proof. By utilising Lemma 2.14 of [**?**], we have that F_∞ is maximal iff for all $m \geq 1$ and all t, $t \in m_w \implies F_\infty(t) \in (m-1)_w$. We proceed by induction on t.

If $t = \mathsf{f}\langle\, \overline{\tau}\, \rangle\ \overline{s'}\ u\ \overline{s_n}$ or $t = x\ \overline{s'}\ u\ \overline{s_n}$ where all members of $\overline{s'}$ are in normal form, $\|u\| > 0$, $u \in m_w^0, s_1 \in m_w^1 \ldots s_n \in m_w^n$ and $m = m^0 + \ldots + m^n$, then $F_\infty(t) = \zeta\ \overline{s'}\ F_\infty(u)\ \overline{s_n}$. By the induction hypothesis $F_\infty(u) \in (m^0 - 1)_w$. Thus, $F_\infty(t) = \zeta\ \overline{s'}\ F_\infty(u)\ \overline{s_n} \in (m-1)_w$.

If $t = \tilde{\mathcal{C}}\ \overline{t_n}$, the proof splits into two:

$\tilde{\mathcal{C}} \neq \mathsf{K}$ or $\|t_1\| = 0$ Then $F_\infty(t) = (\tilde{\mathcal{C}}\ \overline{t_n}\ ){\downarrow}^w$. By the fundamental lemma of maximality, we have $\|(\tilde{\mathcal{C}}\ \overline{t_n}\ ){\downarrow}^w\ \| + 1 = \|\tilde{\mathcal{C}}\ \overline{t_n}\ \| \geq m$. Thus $\|(\tilde{\mathcal{C}}\ \overline{t_n}\ ){\downarrow}^w\ \| \geq m - 1$ and $F_\infty(t) \in (m-1)_n$.
$\tilde{\mathcal{C}} = \mathsf{K}$ and $\|t_1\| > 0$ By the fundamental lemma of maximality we have that $\|(\tilde{\mathcal{C}}\ \overline{t_n}\ ){\downarrow}^w\ \| + \|t_1\| + 1 = \|\tilde{\mathcal{C}}\ \overline{t_n}\ \| \geq m$. By the induction hypothesis we have that $\|F_\infty(t_1)\| \geq \|t_1\| - 1$. Thus

$$
\begin{array}{rr}
\|F_\infty(t)\| = & \|(t){\downarrow}^w\ \| + \|F_\infty(t_1)\| + 1 \\
\geq & \|(t){\downarrow}^w\ \| + \|t_1\| \\
= & \|t\| - 1 \\
\geq & m - 1
\end{array}
$$

and so $F_\infty(t) \in (m-1)_w$.

3 Term Order

First, Blanchette et al.'s [**?**] graceful higher-order basic KBO is presented as it is utilised within our ordering. The presentation here differs slightly from that in [**?**] because we do not allow ordinal weightings and all function symbols have finite arities. Furthermore, we do not allow the use of different operators for the comparison of tuples, but rather restrict the comparison of tuples to use only the length-lexicographic extension of the base order. This is denoted $\gg_{\mathsf{hb}}^{\mathsf{length\_lex}}$. For tuples of terms $\overline{t_n}$ and $\overline{s_n}$ of the same length, the length-lexicographic and lexicographic operators are the same, For this section, terms are assumed to be untyped following the original presentation.

3.1 Graceful Higher-Order Basic KBO

Standard first-order KBO first compares the weights of terms, then compares their head-symbols and finally compares arguments recursively. When working with higher-order terms, the head symbol may be a variable. To allow the comparison of variable heads, a mapping ghd is introduced that maps variable heads to members of Σ that could possibly instantiate the head. This mapping *respects arities* if for any variable x, all members of $ghd(x)$ have arities greater or equal to that of x. The mapping can be extended to constant heads by taking $ghd(\mathsf{f}) = \{\mathsf{f}\}$. A substitution σ *respects* the mapping ghd, if for all variables x, $ghd(x\sigma) \subseteq ghd(x)$.

Let $\succ$ be a total well-founded ordering or *precedence* on Σ. The precedence $\succ$ is extended to arbitrary heads by defining $\zeta \succ \xi$ iff $\forall f \in ghd(\zeta)$ and $\forall g \in ghd(\xi)$, $f \succ g$. Let w be a function from Σ to $\mathbb{N}$ that denotes the weight of a function symbol and $\mathcal{W}$ a function from $\mathcal{T}$ to $\mathbb{N}$ denoting the weight of a term. Let $\varepsilon \in \mathbb{N}_{>0}$. For all constants c, $w(\mathsf{c}) \geq \varepsilon$. The weight of a term is defined recursively:

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

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

Z1 $\mathcal{W}(t) > \mathcal{W}(s)$
Z2 $\mathcal{W}(t) = \mathcal{W}(s)$ and $\zeta \succ \xi$
Z3 $\mathcal{W}(t) = \mathcal{W}(s), \zeta = \xi$ and $\bar{t} \gg_{\mathsf{hb}}^{\mathsf{length\_lex}} \bar{s}$

3.2 Combinator Orienting KBO

The combinator orienting KBO is the focus of this paper. It has the property that all ground instances of combinator axioms are oriented by it left-to-right. This is achieved by first comparing terms by the length of the longest weak reduction from the term and then using $>_{\mathsf{hb}}$. This simple approach runs into problems with regards to stability under substitution, a crucial feature for any ordering used in superposition.

Consider the terms $t = \mathsf{f}\ x\ \mathsf{a}$ and $s = x\ \mathsf{b}$. As the length of the maximum reduction from both terms is 0, the terms would be compared using $>_{\mathsf{hb}}$ resulting in $t \succ s$ as $\mathcal{W}(t) > \mathcal{W}(s)$. Now, consider the substitution $\theta = \{x \to \mathsf{I}\}$. Then, $\|s\theta\| = 1$ whilst $\|t\theta\| = 0$ resulting in $s\theta \succ t\theta$.

To avoid problems such as the above and to ensure that the order enjoys stability under substitution we require some conditions on the forms of terms that can be compared. We provide two approaches. First, in the spirit of Bentkamp et al. [**?**], we provide a translation that replaces "problematic" subterms of the terms to be compared with fresh variables. With this approach, the simple variable condition of the standard KBO, $vars_{\#}(s) \subseteq vars_{\#}(t)$, ensures stability. However, as will be seen, this approach is over constrained and prevents the comparison of terms such as $t = x\ \mathsf{a}$ and $s = x\ \mathsf{b}$ despite the fact that for all substitutions θ, $\|t\theta\| = \|s\theta\|$. Therefore, we also present a second approach wherein no replacement of subterms occurs. This comes at the expense of a far more complex variable condition. Roughly, the condition stipulates that two terms are comparable if and only if the variables and relevant combinators are in identical positions in each.

Approach 1 Because the $>_{\mathsf{hb}}$ ordering is not defined over typed terms, type arguments are replaced by equivalent term arguments before comparison. The translation $(\![\,]\!)$ from $\mathcal{T}$ to untyped terms is given below. First we define precisely the subterms that require replacing by variables.

Definition 1 (Type-1 term). *Consider a term t of the form $\tilde{\mathcal{C}}\ \overline{t_n}$. If there exists a position p such $t|_p$ is a variable, then t is a type-1 term.*

Definition 2 (Type-2 term). *A term $x\ \overline{t_n}$ where $n > 0$ is a type-2 term.*

$$(\![t]\!) = \begin{cases} \tau & t \text{ is a type variable } \tau \\ \kappa\,(\![\,\overline{\sigma_n}\,]\!) & t = \kappa(\,\overline{\sigma_n}\,) \\ x & t \text{ is a term variable } x \\ x_t & t \text{ is a type-1 or type-2 term} \\ \mathsf{f}\,(\![\,\overline{\tau_n}\,]\!) & t = \mathsf{f}\langle\,\overline{\tau_n}\,\rangle \\ (\![t_1]\!)(\![t_2]\!) & t = t_1\,t_2 \end{cases}$$

An untyped term t weak educes to an untyped term t' in one step 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 aim of the ordering presented here is to parametrise the superposition calculus. For this purpose, the property that for terms t and t', $t \longrightarrow_w t' \implies t \succ t'$, is desired. To this end, the following lemma is proved.

Lemma 3. *For all term ground polymorphic terms t and t', it is the case that $t \longrightarrow_w t' \iff (\![t]\!) \longrightarrow_w (\![t']\!)$.*

Proof. The $\implies$ direction is proved by induction on t. If the reduction occurs at ϵ, then t is of the form $\tilde{\mathcal{C}}\langle\,\overline{\tau_n}\,\rangle\ \overline{t_n}$. We prove that the lemma holds if $\tilde{\mathcal{C}} = \mathsf{S}$. The other cases are similar. If $t = \mathsf{S}\langle\tau_1,\tau_2,\tau_3\rangle\ \overline{t_n}$, then $(\![t]\!) = \mathsf{S}\,\tau_1\,\tau_2\,\tau_3\,(\![t_0]\!)(\![t_1]\!)(\![t_2]\!)(\![t_3 \ldots t_n]\!) \longrightarrow_w (\![t_0]\!)(\![t_2]\!)((\![t_1]\!)(\![t_2]\!))(\![t_3 \ldots t_n]\!) = (\![t_0\,t_2\,(t_1\,t_2)\ t_3 \ldots t_n]\!) = (\![t']\!)$. Now assume that the reduction does not occur at ϵ. In this case, $t = \xi\ \overline{t_n}$, $t_i \longrightarrow_w t'_i$ and $t' = \xi\,t_0 \ldots t_{i-1}\,t'_i\,t_{i+1} \ldots t_n$. By the induction hypothesis, $(\![t_i]\!) \longrightarrow_w (\![t'_i]\!)$. Thus, $(\![t]\!) = (\![\xi]\!)(\![\,\overline{t_n}\,]\!) \longrightarrow_w (\![\xi]\!)(\![\,\overline{t_{i-1}}\,]\!)(\![t'_i]\!)(\![t_{i+1} \ldots t_n]\!) = (\![t']\!)$.

The $\Longleftarrow$ direction can be proved in a nearly identical manner.

Corollary 1. *A straightforward corollary of the above Lemma is that for all term-ground polymorphic terms t, $\|t\| = \|(\![t]\!)\|$.*

The combinator orienting Knuth-Bendix order (approach 1) $>_{\mathsf{ski1}}$ is defined as follows. For terms t and s, let $t' = (\![t]\!)$ and $s' = (\![s']\!)$. Then $t >_{\mathsf{ski1}} s$ if $vars_{\#}(s') \subseteq vars_{\#}(t')$ and:

R1 $\|t'\| > \|s'\|$ or,
R2 $\|t'\| = \|s'\|$ and $t' >_{\mathsf{hb}} s'$.

Approach 2 Using approach 1, terms $t = y\,\mathsf{a}$ and $s = y\,\mathsf{b}$ are incomparable. Both are type-2 terms and therefore $(\![t]\!) = x_t$ and $(\![s]\!) = x_s$. The variable condition obviously fails to hold between x_t and x_s. Therefore, we consider another approach which does not replace subterms with fresh variables. We introduce a new translation $[\![]\!]$ from $\mathcal{T}$ to untyped terms that merely replaces type arguments with equivalent term arguments and does not affect term arguments at all. The simpler translation comes at the cost of a more complex variable condition. Before the revised variable definition can be provided, some further terminology requires introduction.

Definition 3 (Safe Combinator). *Let $\tilde{\mathcal{C}}$ occur in t at position p and let p' be the shortest prefix of p such that $head(t|_{p'})$ is a combinator and for all positions p'' between p and p', $head(t|_{p''})$ is a combinator. Let p'' be a prefix of p of length one shorter than p' if such a position exists and ϵ otherwise. Then $\tilde{\mathcal{C}}$ is* safe *in t if $t|_{p'}$ is ground and $head(t|_{p''}) \notin \mathcal{V}$ and* unsafe *otherwise.*

Intuitively, unsafe combinators are those that could affect a variable on a longest reduction path or could become applied to a subterm of a substitution. For example, all combinators in the term $\mathbf{S}\,(\mathbf{K}\,\mathbf{I})\,\mathsf{a}\,x$ are unsafe because they affect x, whilst the combinator in $\mathsf{f}\,(\mathbf{I}\,\mathsf{b})\,y$ is safe. The combinators in $x\,(\mathbf{S}\,\mathbf{I})\,\mathsf{a}$ are unsafe because they could potentially interact with a term substituted for x.

Definition 4. *We say a subterm is* top-level *in a term t if it doesn't appear beneath an applied variable or fully applied combinator head in t.*

In the variable condition, if t is an untyped term of the form $x\,\overline{t_m}$ or $\tilde{\mathcal{C}}\,\overline{t_m}$ and s an untyped term $x\,\overline{s_m}$ or $\tilde{\mathcal{C}}\,\overline{s_m}$, we say that $t \leq s$ ($t \geq s$) if for $0 \leq i \leq m$, $\|t_i\| \leq \|s_i\|$ ($\|t_i\| \geq \|s_i\|$).

Variable Condition:
Let $t' = [\![t]\!]$ and $s' = [\![s]\!]$ for polymorphic terms t and s. Let A be the multiset of all top-level, non-ground, first-order subterms in s' of the form $x\,\overline{s_n}$ (n may be 0) or $\tilde{\mathcal{C}}\,\overline{t_n}$. Let B be a similarly defined multiset of subterms over t'. Then, $var\_cond(t', s')$ holds if there exists an injective total function f from A to B such that f only associates terms t_1 and t_2 if for all p in $pos(t_1)$ such that $t_1|_p = \tilde{\mathcal{C}}\,\overline{t_n}$ and $\tilde{\mathcal{C}}$ (not necessarily fully applied) is unsafe, then $t_2|_p = \tilde{\mathcal{C}}\,\overline{s_n}$ and $t_2|_p \geq t_1|_p$. Further, for all p in $pos(t_1)$ such that $t_1|_p = x\,\overline{t_n}$, then $t_2|_p = x\,\overline{s_n}$ and $t_2|_p \geq t_1|_p$.

For example $var\_cond(t, s)$ holds where $t = \mathsf{f}\,y\,(x\,\mathsf{a})$ and $s = \mathsf{g}\,(x\,b)$. In this case $A = \{x\,\mathsf{b}\}$ and $B = \{y, x\,\mathsf{a}\}$. There exists and injective total function from A to B that matches the requirements by relating $x\,\mathsf{b}$ to $x\,\mathsf{a}$. However, the variable condition does not hold in either direction if $t = \mathsf{f}\,y\,(x\,\mathsf{a})$ and $s = \mathsf{g}\,(x\,(\mathbf{I}\,\mathsf{b}))$. In this case, $x\,(\mathbf{I}\,\mathsf{b})$ cannot be related to $x\,a$ since the condition that $x\,a \geq x\,(\mathbf{I}\,\mathsf{b})$ is not fulfilled.

The combinator orienting Knuth-Bendix order (approach 2) $>_{\mathsf{ski}}$ is defined as follows. For terms t and s, let $t' = [\![t]\!]$ and $s' = [\![s']\!]$. Then $t >_{\mathsf{ski}} s$ if $var\_cond(t', s')$ and:

R1 $\|t'\| > \|s'\|$ or,
R2 $\|t'\| = \|s'\|$ and $t' >_{\mathsf{hb}} s'$.

Lemma 4. *For all ground instances of the combinator axioms $l \approx r$, it is the case that $l >_{\mathsf{ski}} r$.*

Proof. Since for all ground 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. . The proofs can easily be modified to hold for the less powerful $>_{\mathsf{ski1}}$ ordering. Note that the $>_{\mathsf{ski}}$ order is not a reduction ordering as it lacks full compatibility with contexts.

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

Proof. Let $s' = [\![s]\!]$. 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 $>_{\mathsf{hb}}$.

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. Let $s' = [\![s]\!]$, $t' = [\![t]\!]$ and $u' = [\![u]\!]$. First we prove that if $var\_cond(t', s')$ and $var\_cond(s', u')$, then $var\_cond(t', u')$. Since $var\_cond(s', u')$ holds, there exists a function f_1 from the the multiset of top-level, non-ground, first-order subterms in u' to the like multiset in s' that meets the given requirements. There is a similar function f_2 for s' and t'. We show that $f_2 \circ f_1$ is a function with the desired characteristics for terms u' and t'. Since f_1 and f_2 are both injective and total, $f_2 \circ f_1$ must be injective and total. Now suppose that f_1 relates a subterm of u', u_1 to a subterm of s', s_1 and f_2 relates s_1 to a subterm of t', t_1. Let p be any position in u_1 such $u_1|_p = \tilde{\mathcal{C}}\ \overline{u_m}$ where $\tilde{\mathcal{C}}$ is unsafe. By the variable condition, it must be the case that $s_1|_p = \tilde{\mathcal{C}}\ \overline{s_m}$ and $s_1|_p \geq u_1|_p$. But then $\tilde{\mathcal{C}}$ must be unsafe in s_1 and therefore by the variable condition $t_1|_p = \tilde{\mathcal{C}}\ \overline{t_m}$ and $t_1|_p \geq s_1|_p$. By the transitivity of $\geq$, $t_1|_p \geq u_1|_p$ follows.

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 $s' >_{\mathsf{hb}} t'$ and $t' >_{\mathsf{hb}} u'$. It follows from the transitivity of $>_{\mathsf{hb}}$ that $s' >_{\mathsf{hb}} u'$ 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. Let $s' = [\![s]\!]$ and $t' = [\![t]\!]$. If $\|s'\| \neq \|t'\|$ then by R1 either $s >_{\mathsf{ski}} t$ or $t >_{\mathsf{ski}} s$. Otherwise, s' and t' are compared using $>_{\mathsf{hb}}$ and either $t' >_{\mathsf{hb}} s'$ or $s' >_{\mathsf{hb}} t'$ holds by the ground totality of $>_{\mathsf{hb}}$ and the injectivity of $[\![]\!]$.

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

Proof. Let $s' = [\![s]\!]$ and $t' = [\![t]\!]$. Since t is a subterm of s, t' is a subterm of s' and $\|s'\| \geq \|t'\|$ because any weak reduction in t' is also a weak reduction in s'. If $\|s'\| > \|t'\|$, the theorem follows by an application of R1. Otherwise s' and t' are compared using $>_{\mathsf{hb}}$ and $s' >_{\mathsf{hb}} t'$ holds by the subterm property of $>_{\mathsf{hb}}$. Thus $s >_{\mathsf{ski}} t$.

Next, a series of Lemmas are proved that are utilised in the proof of the ordering's compatability with contexts and stability under substitution. We prove two monotonicity properties Lemma 9 and Theorem 5. Both hold for non-ground terms, but to show this, it is required to show that the variable condition holds between terms $u[t]$ and $u[s]$ for t and s such that $t >_{\mathsf{ski}} s$. To avoid this complication, we prove the Lemmas for ground terms which suffices for our purposes. To avoid clutter, assume that the terms mentioned in the statement of Lemmas 5, 6, 7, 8, 10, 11, 14, 12 and 13 are all untyped terms formed by translating polymorphic terms.

Lemma 5. $\|\zeta\,\overline{t_n}\,\| = \sum_{i=1}^{n} \|t_i\|$ *if ζ is not a fully applied combinator.*

Proof. Trivial.

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. 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 $s = \zeta\ \overline{s_n}$ and $s' = \zeta\ \overline{s'_n}$ where ζ is not a fully applied combinator other than K.

Assume that $\zeta \neq \mathsf{K}$. From the definition of F_∞, we have $F_\infty(s) = \zeta\, s_1 \ldots s_{i-1}$ $F_\infty(s_i)\, s_{i+1} \ldots s_n$ where $\|s_j\| = 0$ for $1 \leq j < i$. Since for all i, s_i and s'_i only differ at positions where one contains a t_j and the other contains a t'_j and $\|t_i\| \geq \|t'_i\|$ for $1 \leq i \leq n$, we have that $\|s_j\| = 0$ implies $\|s'_j\| = 0$. Thus, using the definition of F_∞, $F_\infty(s') = \zeta\, s'_1 \ldots s'_{i-1}\, F_\infty(s'_i)\, s'_{i+1} \ldots s'_n$. 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 unless $\|t'_i\| = 0$. The lemma follows immediately.

Assume that $\zeta = \mathsf{K}$. It must be the case that $\|s_2\| > 0$ otherwise the F_∞ redex in s would be at the head and not within a t_i. By the definition of F_∞, $F_\infty(s) = \mathsf{K}\, s_1\, F_\infty(s_2)\, s_3 \ldots s_n$. Let the F_∞ redex of s_2 occur inside t_j. Then t'_j is a subterm of s'_2. If $\|t'_j\| > 0$ then $\|s'_2\| > 0$ and $F_\infty(s') = \mathsf{K}\, s'_1\, F_\infty(s'_2)\, s'_3 \ldots s'_n$. By the induction hypothesis, the F_∞ redex of s'_2 occurs in t'_j.

Lemma 7. *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 $u[\,\overline{t_n}\,] \longrightarrow_w u'$ 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 $p > p_i$ for some i, there exists a $p' \neq \epsilon$ such that $p = p_i p'$. Then, $u[t_1, \ldots, t_i, \ldots, t_n]|_{p_i} = t_i[\tilde{\mathcal{C}}\,\overline{r_n}\,]_{p'} \longrightarrow_w t_i[(\tilde{\mathcal{C}}\,\overline{r_n}\,)\downarrow^w]_{p'}$. Let $\hat{t_i} = t_i[(\tilde{\mathcal{C}}\,\overline{r_n}\,)\downarrow^w]_{p'}$. We thus have that $t_i \longrightarrow_w \hat{t_i}$ and thus $u[t_1, \ldots, t_i, \ldots, t_n] \longrightarrow_w u[t_1, \ldots, \hat{t_i}, \ldots, t_n]$.

It cannot be the case that $p = p_i$ for any i because $head(t_i)$ is not a combinator for any t_i. In the case where $p < p_i$ or $p \parallel p_i$ for all i, we have that $u[\,\overline{t_n}\,] = (u[\,\overline{x_n}\,])\sigma$ and $u[\,\overline{x_n}\,]|_p$ is a redex where $\sigma = \{\,\overline{x_n} \to \overline{t_n}\,\}$. Let $\hat{u}$ be formed from $u[\,\overline{x_n}\,]$ by reducing its redex at p. Then:

$$
\begin{aligned}
s = u[\,\overline{t_n}\,] &= && (u[\,\overline{x_n}\,])\sigma \\
&\longrightarrow_w && \hat{u}\sigma \\
&= && \hat{u}\{x_1 \to t_1 \ldots x_n \to t_n\}
\end{aligned}
$$

Lemma 8. *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* $\|u[\,\overline{t_n}\,]\| = \|u[\,\overline{t'_n}\,]\|$ *for all* n *holed contexts* u.
2. *If* $\|t_j\| > \|t'_j\|$ *for some* $j \in \{1, \ldots, n\}$ *and* $\|t_i\| \geq \|t'_i\|$ *for* $i \neq j$, *then* $\|u[\,\overline{t_n}\,]\| > \|u[\,\overline{t'_n}\,]\|$ *for all* n *holed contexts* u.

Proof. 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 $s' = u[\,\overline{t'_n}\,]$. Proof is by induction on $\|s\| + \|s'\|$. 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 $\|u[\,\overline{t'_n}\,]\| = 0$.

Assume that $\|u[\,\overline{t_n}\,]\| > 0$. Let $u' = F_\infty(s)$. By Lemma 7 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 $u' = \hat{u}\{\,\overline{x_n} \to \overline{t_n}\,\}$ where $u[\,\overline{x_n}\,] \longrightarrow_w \hat{u}$. In the first case, by Lemma 6 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, $F_\infty(s') = u'' = \hat{u}\{\,\overline{x_n} \to \overline{t'_n}\,\}$ 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 6 $u'' = u[t'_1, \ldots, \hat{t'_i}, \ldots, t'_n]$ where $t'_i \longrightarrow_w \hat{t'_i}$ unless $\|t'_i\| = 0$ and 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 $u' = \hat{u}\{\,\overline{x_n} \to \overline{t_n}\,\}$, then $u'' = \hat{u}\{\,\overline{x_n} \to \overline{t'_n}\,\}$ and the lemma follows by the induction hypothesis.

Lemma 9 (Compatibility with Contexts). *For ground 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* $u[s] >_{\mathsf{ski}} u[t]$ *for all ground contexts* $u[]$.

Proof. Let $s' = [\![s]\!]$, $t' = [\![t]\!]$ and $u' = [\![u]\!]$. By Lemma 8 Part 2, 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 8 Part 1, $\|u'[s']\| = \|u'[t']\|$ follows. Thus, $u'[s']$ is compared with $u'[t']$ by R2 and $u[s] >_{\mathsf{ski}} u[t]$ by the compatibility with contexts of $>_{\mathsf{hb}}$.

Corollary 2 (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. This is just a special case of Lemma 9.

Lemma 10. $\|s\| > \|t\| \implies \|u\langle\!\langle s\rangle\!\rangle\| > \|u\langle\!\langle t\rangle\!\rangle\|$ *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 the first implication of the lemma first. If u is not the empty context, $u\langle\!\langle s\rangle\!\rangle$ is of the form $u'\langle\!\langle \zeta\, t_1 \dots t_{i-1}, s, t_{i+1} \dots t_n\rangle\!\rangle$. By the definition of a stable subterm ζ cannot be a fully applied combinator and thus by Lemma 5 we have:

$$\begin{aligned}
\|\zeta\, t_1 \dots t_{i-1}, s, t_{i+1} \dots t_n\| &= \sum_{\substack{j=1\\ j\neq i}}^{n} \|t_j\| + \|s\| \\
&> \sum_{\substack{j=1\\ j\neq i}}^{n} \|t_j\| + \|t\| \\
&= \|\zeta\, t_1 \dots t_{i-1}, t, t_{i+1} \dots t_n\|
\end{aligned}$$

If ζ is not a combinator, then $\|u'\langle\!\langle \zeta\, t_1 \dots t_{i-1}, s, t_{i+1} \dots t_n\rangle\!\rangle\| > \|u'\langle\!\langle \zeta\ t_1 \dots t_{i-1}, t, t_{i+1} \dots t_n\rangle\!\rangle\|$ follows from Lemma 8 Part 2. Otherwise, as u' is smaller stable context than u, the induction hypothesis can be used to conclude that $\|u'\langle\!\langle \zeta\, t_1\ \dots t_{i-1}, s, t_{i+1} \dots t_n\rangle\!\rangle\| > \|u'\langle\!\langle \zeta\, t_1 \dots t_{i-1}, t, t_{i+1} \dots t_n\rangle\!\rangle\|$ and thus that $\|u\langle\!\langle s\rangle\!\rangle\| > \|u\langle\!\langle t\rangle\!\rangle\|$. The proof of the inductive case for the second implication of the lemma is almost identical.

Theorem 5 (Compatibility with Stable Contexts). *For all stable ground contexts* $u\langle\!\langle\rangle\!\rangle$ *and ground terms s and t, 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 10, $\|u\langle\!\langle s\rangle\!\rangle\| > \|u\langle\!\langle t\rangle\!\rangle\|$ 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 10 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 $>_{\mathsf{hb}}$. By the compatibility with contexts of $>_{\mathsf{hb}}$, $[\![u\langle\!\langle s\rangle\!\rangle]\!] >_{\mathsf{hb}} [\![u\langle\!\langle t\rangle\!\rangle]\!]$ holds and then by an application of R2 $u\langle\!\langle s\rangle\!\rangle >_{\mathsf{ski}} u\langle\!\langle t\rangle\!\rangle$ is true.

Lemma 11. *For a single hole context* $u\langle\!\langle\rangle\!\rangle$ *such that the hole does not occur below a fully applied combinator and any term t,* $\|u\langle\!\langle t\rangle\!\rangle\| = \|u\langle\!\langle\rangle\!\rangle\| + \|t\|$.

Proof. Proceed by induction on the size of u. If u is the empty context the theorem follows trivially. Therefore, assume that $u = \mathsf{f}\, t \ldots t_{i-1}\, u'\langle\langle\rangle\rangle\, t_{i+1} \ldots t_n$. By Lemma 5 $\|u\langle\langle\rangle\rangle\| = \sum_{i=1}^{n} \|t_i\| + \|u'\langle\langle\rangle\rangle\|$. Because u' is a smaller context than u, the induction hypothesis can be used to show $\|u'\langle\langle t\rangle\rangle\| = \|u'\langle\langle\rangle\rangle\| + \|t\|$. Thus:

$$\begin{aligned}
\|u\langle\langle t\rangle\rangle\| &= \sum_{i=1}^{n} \|t_i\| + \|u'\langle\langle t\rangle\rangle\| \\
&= \sum_{i=1}^{n} \|t_i\| + \|u'\langle\langle\rangle\rangle\| + \|t\| \\
&= \|u\langle\langle\rangle\rangle\| + \|t\|
\end{aligned}$$

proving the theorem.

Lemma 12. *Let $\overline{t_n}$ and $\overline{s_n}$ be terms such that for $n_i \ldots n_n \in \mathbb{N}$ and for $0 \leq i \leq n$, $\|t_i\| \geq \|s_i\| + n_i$. Further, let $t = t_0\, t_1 \ldots t_n$ and $s = s_0\, s_1 \ldots s_n$. Assume that neither t nor s is ground and that $var\_cond(t, s)$ holds. Then $\|t\| \geq \|s\| + \sum_{i=0}^{n} n_i$.*

Proof. If $head(t)$ is not a fully applied combinator then by the variable condition $head(s)$ cannot be a fully applied combinator. In this case, $\|t\| = \|t_0\| + \ldots + \|t_n\| \geq \|s_0\| + \ldots + \|s_n\| + n_0 + \ldots + n_n$ and the Lemma follows.

If $head(t)$ is a fully applied combinator, we proceed by induction on $\|t\| + \|s\|$. The proof splits into two depending on whether $head(t_0)$ (and therefore $head(s_0)$) is a fully applied combinator or not.

If $t_0 = \tilde{\mathcal{C}}\ \overline{t'_m}$, the variable condition ensures that $s_0 = \tilde{\mathcal{C}}\ \overline{s'_m}$. Let $t' = F_\infty(t_0)$ and $s' = F_\infty(s_0)$. Then, $\|t_0\| = \|t'\| + 1$ and $\|s_0\| = \|s'\| + 1$, so $\|t'\| \geq \|s'\| + n_0$. Since $\|t'\, t_1 \ldots t_n\| + \|s'\, s_1 \ldots s_n\| < \|t\| + \|s\|$, the induction hypothesis can be invoked to conclude that $\|t'\, t_1 \ldots t_n\| \geq \|s'\, s_1 \ldots s_n\| + \sum_{i=0}^{n} n_i$. From this, $\|t\| \geq \|s\| + \sum_{i=0}^{n} n_i$ follows since $\|t\| = 1 + \|t'\, t_1 \ldots t_n\|$ and $\|s\| = 1 + \|s'\, s_1 \ldots s_n\|$.

If $head(t_0)$ is a partially applied combinator, the proof follows by a case analysis of the particular combinator along with its argument number. We show a couple of cases and the remainder can be proved in a similar fashion.

Assume that $t_0 = \mathsf{I}$. Then $s_0 = \mathsf{I}$ as well and $n_1 = 0$. For terms $t' = F_\infty(t) = t_1 \ldots t_n$ and $s' = F_\infty(s) = s_1 \ldots s_n$, the induction hypothesis gives that $\|t'\| \geq \|s'\| + \sum_{i=1}^{n} n_i$. Then $\|t\| \geq \|s\| + \sum_{i=1}^{n} n_i + 0 = \|s\| + \sum_{i=0}^{n} n_i$.

Assume that $t_0 = \mathsf{S}\, t'_0\, t'_1$. Then $s_0 = \mathsf{S}\, s'_0\, s'_1$. Further, $\|t'_0\| \geq \|s'_0\| + n'_0$ and $\|t'_1\| \geq \|s'_1\| + n'_1$ where $n'_0 + n'_1 = n_0$. Let $t' = F_\infty(t) = t'_0\, t_1\, (t'_1\, t_1)\, t_2 \ldots t_n$ and $s' = F_\infty(s) = s'_0\, s_1\, (s'_1\, s_1)\, s_2 \ldots s_n$. The induction hypothesis can be utilised on $t'_1\, t_1$ and $s'_1\, s_1$ to give $\|t'_1\, t_1\| \geq \|s'_1\, s_1\| + n'_1 + n_1$. The induction hypothesis can be used a second time on t' and s' to give $\|t'\| \geq \|s'\| + n'_0 + n_1 + n'_1 + n_1 + n_2 + \ldots + n_n$. Since $n'_0 + n_1 + n'_1 + n_1 + n_2 + \ldots + n_n \geq \sum_{i=0}^{n} n_i$, it can be concluded that $\|t\| \geq \|s\| + \sum_{i=0}^{n} n_i$.

Lemma 13. *For terms t and s such that $var\_cond(t, s)$ holds and $\|t\| \geq \|s\| + n$ for some $n \in \mathbb{N}$, for all substitutions θ, $\|t\theta\| \geq \|s\theta\| + n$.*

Proof. Proof by induction on the pair $(\|t\| + \|s\|, |t| + |s|)$. Pairs are compared lexicographically left-to-right. If s and t are ground, the theorem is trivial. If s is ground, then $\|t\theta\| > \|t\| \geq \|s\| + n$. If s is not ground, then $var\_cond(t, s)$ implies that t is not ground. The proof of the lemma for non-ground s and t is broken into three cases.

1. $\boldsymbol{head(s), head(t)}$ are fully applied combinators: by the variable condition, s and t must have the same head, and therefore $t = \tilde{\mathcal{C}}\ \overline{t_n}$ and $s = \tilde{\mathcal{C}}\ \overline{s_n}$. If $\tilde{\mathcal{C}} \neq \mathsf{K}$, consider the terms $t' = (t)\downarrow^w$ and $s' = (s)\downarrow^w$. We have that, $var\_cond(t', s')$ holds. Furthermore, since $\|t\| = \|t'\| + 1$ and $\|s\| = \|s'\| + 1$, $\|t'\| \geq \|s'\| + n$ holds as well. As $\|t'\| + \|s'\| < \|t\| + \|s\|$, the induction hypothesis can be applied to show that $\|t'\theta\| \geq \|s'\theta\| + n$. Then $\|t\theta\| \geq \|s\theta\| + n$ follows since $\|t\theta\| = \|t'\theta\| + 1$ and $\|s\theta\| = \|s'\theta\| + 1$.

On the other hand, if $\tilde{\mathcal{C}} = \mathsf{K}$, then $\|t\| = \|t_0\, t_2 \ldots t_n\| + \|t_1\| + 1$ by the fundamental lemma of maximality. Similarly $\|s\| = \|s_0\, s_2 \ldots s_n\| + \|s_1\| + 1$. Further $\|t_1\| \geq \|s_1\| + n_1$ and $\|t_0\, t_2 \ldots t_n\| \geq \|s_0\, s_2 \ldots s_n\| + n_2$ where $n_1 + n_2 = n$. Since $\|s_1\| + \|t_1\| < \|s\| + \|t\|$, the induction hypothesis can be invoked to conclude that $\|t_1\theta\| \geq \|s_1\theta\| + n_1$. Likewise, the induction hypothesis gives $\|(t_0\, t_2 \ldots t_n)\theta\| \geq \|(s_0\, s_2 \ldots s_n)\theta\| + n_2$. This case of the proof can be concluded by observing:

$$\begin{aligned}
\|t\theta\| &= \|(\mathsf{K}\ \overline{t_n}\,)\theta\| \\
&= \|(t_0\, t_2 \ldots t_n)\theta\| + \|t_1\theta\| + 1 \\
&\geq \|(s_0\, s_2 \ldots s_n)\theta\| + n_2 + \|s_1\theta\| + n_1 + 1 \\
&= \|s\theta\| + n
\end{aligned}$$

2. $\boldsymbol{head(s), head(t)}$ are variables: by the variable condition, $t = x\ \overline{t_n}$ and $s = x\ \overline{s_n}$. If $head(t\theta)$ is not a fully applied combinator, the proof is straightforward. In this case, $t\theta = \xi\, \overline{r_m}\ t_0\theta \ldots t_n\theta$ and $s\theta = \xi\, \overline{r_m}\ s_0\theta \ldots s_n\theta$. For $0 \leq i \leq n$, we have that $\|s_i\| + \|t_i\| \leq \|s\| + \|t\|$. Further, $|s_i| + |t_i| < |s| + |t|$. From $var\_cond(t, s)$, follows $\|t_i\| \geq \|s_i\| + n_i$ and $var\_cond(t_i, s_i)$ where $\sum_{i=1}^{n} n_i = n$. Therefore $\|t_i\theta\| \geq \|s_i\theta\| + n_i$ follows from the induction hypothesis. To show that this implies $\|t\theta\| \geq \|s\theta\| + n$:

$$\begin{aligned}
\|t\theta\| &= \|\xi\, \overline{r_m}\ t_0\theta \ldots t_n\theta\| \\
&= \sum_{i=1}^{m} \|r_i\| + \sum_{i=1}^{n} \|t_i\theta\| \\
&\geq \sum_{i=1}^{m} \|r_i\| + \sum_{i=1}^{n} \|s_i\theta\| + n_i \\
&= \|s\theta\| + n
\end{aligned}$$

If $head(t\theta)$ is a fully applied combinator, then $t\theta = \tilde{\mathcal{C}}\ \overline{r_m}\ \ \overline{t_n\theta}$ and $s\theta = \tilde{\mathcal{C}}\ \overline{r_m}\ \ \overline{s_n\theta}\,$. An n fold application of the induction hypothesis gives $\|t_i\theta\| \geq \|s_i\theta\| + n_i$ for $1 \leq i \leq n$. It is obvious that $var\_cond(t\theta, s\theta)$ holds and therefore Lemma 12 can be used to conclude that $\|t\theta\| \geq \|s\theta\| + 0 + \sum_{i=0}^{n} n_i = \|s\theta\| + n$.

3. $head(s), head(t)$ are not variables nor fully applied combinators: Let $t = u\langle\!\langle\, \overline{t_n}\, \rangle\!\rangle$ and $s = u'\langle\!\langle\, \overline{s_m}\, \rangle\!\rangle$ where $\overline{t_n}$ and $\overline{s_m}$ are all the non-ground top-level first-order subterms of the form $x\ \overline{args}$ or $\tilde{\mathcal{C}}\ \overline{args}$ in t and s respectively. By the variable condition, we have that $n \geq m$ and that there exists a total injective function from the s_i to the t_i that respects the given conditions. Without loss of generality, assume that that this function relates s_0 to t_0, s_1 to t_1 and so on. For $0 \leq i \leq m$, $\|t_i\| = \|s_i\| + n_i$ for $n_i \in \mathbb{N}$. If, for some i, t_i and s_i have variable heads, this is straightforward. Otherwise, it follows from Lemma 12. Let $n' = \|u\langle\!\langle\rangle\!\rangle\| - \|u'\langle\!\langle\rangle\!\rangle\|$. Note that n' could be negative. By Lemma 11 $\|t\| = \|u\langle\!\langle\rangle\!\rangle\| + \sum_{i=1}^{n} \|t_i\|$ and $\|s\| = \|u'\langle\!\langle\rangle\!\rangle\| + \sum_{i=1}^{m} \|s_i\|$. Thus, $\|t\| = \|s\| + n' + \sum_{i=1}^{m} n_i$. Therefore, $n' + \sum_{i=1}^{m} n_i \geq n$. The induction hypothesis can be used to show that for all i, $\|t_i\theta\| \geq \|s_i\theta\| + n_i$. Because $u'\langle\!\langle\rangle\!\rangle$ is ground, it follows $\|u\theta\langle\!\langle\rangle\!\rangle\| - \|u'\theta\langle\!\langle\rangle\!\rangle\| \geq n'$. To conclude this case:

$$
\begin{aligned}
\|t\theta\| &= \|u\theta\langle\!\langle\, \overline{t_n\theta}\, \rangle\!\rangle\| \\
&= \|u\theta\langle\!\langle\rangle\!\rangle\| + \sum_{i=1}^{n} \|t_i\theta\| \\
&\geq \|u'\theta\langle\!\langle\rangle\!\rangle\| + \sum_{i=1}^{m} \|s_i\theta\| + n' + \sum_{i=1}^{m} n_i \\
&\geq \|u'\theta\langle\!\langle\rangle\!\rangle\| + \sum_{i=1}^{m} \|s_i\theta\| + n \\
&= \|s\theta\| + n
\end{aligned}
$$

Lemma 14. *Let t be a polymorphic term and σ be a substitution. We define a new substitution ρ such that the domain of ρ is $dom(\sigma)$. Define $y\rho = [\![y\sigma]\!]$. For all terms t, $[\![t\sigma]\!] = [\![t]\!]\rho$.*

Proof. Via a straightforward induction on t.

Theorem 6 (Stability under Substitution). *If $s >_{\mathsf{ski}} t$ then $s\sigma >_{\mathsf{ski}} t\sigma$ for all substitutions σ that respect the ghd mapping.*

Proof. Let $s' = [\![s]\!]$ and $t' = [\![t]\!]$. Let ρ be defined as per Lemma 14. First, we show that if R1 was used to derive $s >_{\mathsf{ski}} t$ and thus $\|s'\| > \|t'\|$ then $\|s'\rho\| > \|t'\rho\|$ and thus $s\sigma >_{\mathsf{ski}} t\sigma$ because $[\![s\sigma]\!] = s'\rho$ and $[\![t\sigma]\!] = t'\rho$.

First we show that if $var\_cond(s', t')$ holds then $var\_cond\ (s'\rho, t'\rho)$ holds. Assume that $var\_cond\ (s'\rho, t'\rho)$ does not hold. This could be because $s'\rho$ contains a position p such that $s'\rho|_p$ is an unsafe combinator $\tilde{\mathcal{C}}$, and $t'\rho|_p$ is not the same combinator. The position p must be beneath a variable in s' otherwise $var\_cond(s', t')$ would fail to hold. Assume that it occurs beneath a variable y that occurs in s' at p'. But then, y must occur in t' at p' because $var\_cond(s', t')$ and thus $t\rho|_p = \tilde{\mathcal{C}}$ contradicting the assumption. Other causes of failure can be shown to lead to contradictions in the same manner.

Furthermore, if $\|s'\| > \|t'\|$, then by Lemma 13 $\|s'\rho\| > \|t'\rho\|$ and $s\sigma >_{\mathsf{ski}} t\sigma$ by an application of R1.

On the other hand, if $\|s'\| = \|t'\|$, then R2 was used to derive $s >_{\mathsf{ski}} t$. By Lemma 13 $\|s'\rho\| \geq \|t'\rho\|$. If $\|s'\rho\| > \|t'\rho\|$, then this is the same as the former case. Otherwise $\|s'\rho\| = \|t'\rho\|$ and $s'\rho$ and $t\rho$ are compared using R2. From the stability under substitution of $>_{\mathsf{hb}}$, $s'\rho >_{\mathsf{hb}} t'\rho$ follows and $s\sigma >_{\mathsf{ski}} t\sigma$ can be concluded.

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 $>_{\mathsf{hb}}$ comparisons $[\![s_m]\!] >_{\mathsf{hb}} [\![s_{m+1}]\!] >_{\mathsf{hb}} [\![s_{m+2}]\!] \cdots$. This contradicts the well-foundedness of $>_{\mathsf{hb}}$.

Theorem 8 (Coincidence with First-Order KBO). *Let $>_{\mathsf{fo}}$ be the first-oder KBO as described by Blanchette et al. in [?]. Assume that $>_{\mathsf{ski}}$ and $>_{\mathsf{fo}}$ are parameterised by the same precedence $\succ$ and that $>_{\mathsf{fo}}$ always compares tuples using the lexicographic extension operator. Then $>_{\mathsf{ski}}$ and $>_{\mathsf{fo}}$ always agree on first-order terms.*

Proof. Let $t' = [\![t]\!]$ and $s' = [\![s]\!]$. Since s and t are first-order, $\|s'\| = 0$ and $\|t'\| = 0$. Thus, s' and t' will always be compared by $>_{\mathsf{hb}}$. Since $>_{\mathsf{hb}}$ coincides with $>_{\mathsf{fo}}$ on first-order terms, so does $>_{\mathsf{ski}}$.

5 Examples

To give a flavour of how the ordering behaves, we provide a number of examples.

Example 1. Consider the terms (ignoring type arguments) $t = \mathbf{S}\,(\mathbf{K}\,\mathsf{a})\,\mathsf{b}\,\mathsf{c}$ and $s = \mathsf{f}\,\mathsf{c}\,\mathsf{e}$. From the definition of the translation $[\![]\!]$, we have that $[\![t]\!] = \mathbf{S}\,(\mathbf{K}\,\mathsf{a})\,\mathsf{b}\,\mathsf{c}$ and $[\![s]\!] = \mathsf{f}\,\mathsf{c}\,\mathsf{e}$. Since $\|\mathbf{S}\,(\mathbf{K}\,\mathsf{a})\,\mathsf{b}\,\mathsf{c}\| = 2$ and $\|\mathsf{f}\,\mathsf{c}\,\mathsf{e}\| = 0$, we have that $t >_{\mathsf{ski}} s$.

Example 2. Consider the terms $t = \mathsf{f}\,(\mathsf{g}\,\mathsf{b})\,\mathsf{e}\,\mathsf{d}$ and $s = \mathbf{I}\,\mathsf{a}$. Here $s >_{\mathsf{ski}} t$ despite the fact that s is syntactically smaller than t because s has a maximum reduction of 1 as opposed to 0 of t.

Example 3. Consider terms $t = \mathsf{f}\,(\mathbf{I}\,\mathsf{d})\,(\mathbf{S}\,x\,\mathsf{a}\,\mathsf{b})$ and $s = \mathsf{g}\,(\mathbf{S}\,x\,(\mathsf{h}\,\mathsf{d})\,\mathsf{b})$. The two terms are comparable as the variable condition relates subterm $\mathbf{S}\,x\,(\mathsf{h}\,\mathsf{d})\,\mathsf{b}$ in s to subterm $\mathbf{S}\,x\,\mathsf{a}\,\mathsf{b}$ in t. The unsafe combinator $\mathbf{S}$ and variable x are in the same position in each subterm. As $\|t\| > \|s\|$, $t >_{\mathsf{ski}} s$.

Example 4. Consider terms $t = \mathsf{f}\,(\mathbf{I}\,\mathsf{d})\,(\mathbf{S}\,x\,\mathsf{a}\,y)$ and $s = \mathsf{g}\,(\mathbf{S}\,x\,(\mathsf{h}\,y)\,\mathsf{b})$. This is very similar to the previous example, but in this case the terms are incomparable. Let s' be a name for the subterm $(\mathbf{S}\,x\,(\mathsf{h}\,y)\,\mathsf{b})$ in s and t' a name for the subterm $(\mathbf{S}\,x\,\mathsf{a}\,y)$. The variable y occurs in different positions in s' and t'. Therefore, s' cannot be related to t by the variable condition and the two terms are incomparable.

Example 5. Consider terms $t = \mathsf{f}\,(x\,(\mathsf{g}\,(\mathbf{K}\,\mathbf{I}\,\mathsf{a}\,\mathsf{b})))$ and $s = \mathsf{h}\,(\mathbf{I}\,\mathsf{a})\,(x\,\mathsf{c})$. The variable condition holds between t and s by relating $(x\,(\mathsf{g}\,(\mathbf{K}\,\mathbf{I}\,\mathsf{a}\,\mathsf{b})))$ to $(x\,\mathsf{c})$. The combinator $\mathbf{I}$ in s is not unsafe and therefore does not need to be related to a combinator in t.

Since $\|t\| = 2 > \|s\| = 1$, $t >_{\mathsf{ski}} s$. Intuitively, this is safe because a susbtituion for x in t can duplicate $(\mathsf{g}\,(\mathbf{K}\,\mathbf{I}\,\mathsf{a}\,\mathsf{b}))$ whose maximum reduction length is 2 whilst a substitution for x in s can only duplicate c whose maximum reduction length is 0.

6 Extending to β-Reduction

The ordering that has been explored above orients all ground instances of combinator equations left-to-right and as a result has the property that for ground terms t_1 and t_2 if $t_1 \longrightarrow_w^+ t_2$ then $t_1 >_{\mathsf{ski}} t_2$. It would be useful if an ordering with a similar property with respect to β-reduction could be developed and this is what is explored in this section. We first define some terminology that will be used throughout the section, then present an ordering analogous to the $>_{\mathsf{ski}}$ ordering.

Polymorphic types and type declarations are as defined previously. Below the set of *raw λ-terms* is defined.

$$\textbf{Terms} \quad \mathcal{T} ::= x \mid \mathsf{f}\langle\,\overline{\tau_n}\,\rangle \mid \lambda x.t \mid t_{1_{\tau_1 \to \tau_2}} t_{2_{\tau_1}}$$
$$\text{where } x \in \mathcal{V},\, t, t_1, t_2 \in \mathcal{T},\, \mathsf{f} \in \Sigma,\, \mathsf{f} : \Pi\,\overline{\alpha_n}\,.\sigma \text{ and } \overline{\tau_n} \text{ are types}$$

If x is of type τ and t is of type σ then the type of $\lambda x.t$ is $\tau \to \sigma$. The type of the term $\mathsf{f}\langle\,\overline{\tau_n}\,\rangle$ is $\sigma\{\,\overline{\alpha_n} \to \overline{\tau_n}\,\}$. Variables that are bound by the λ binder are known as *bound* variables and all other variables are *free* variables. A term that contains no type variables or free term variables is known as a *ground* term. The consistent renaming of a bound variable throughout a term is known as *α-renaming*. For example the term $\lambda x.\mathsf{f}\,x$ can be α-renamed to $\lambda y.\mathsf{f}\,y$. Raw λ-terms can be partitioned into equivalence classes modulo α-renaming. These equivalence classes are known as λ-terms.

Positions over λ-terms are defined similarly to positions over combinatory terms with the added definition $pos(\lambda x.t) = \{\epsilon\} \cup \{1.p \mid p \in pos(t)\}$. First-order subterms are defined exactly as before. However, stable subterms require redefining. Again, stable subterms are a subset of first-order subterms. By definition, if $p.2 \in pos(t)$ for some term t, then the subterm at $p.2$ is first-order. The subterm is *stable* if $head(t|_p)$ is not a λ-expression or $head(t|_p) = \lambda\,\overline{y_n}\,.t'$, $head(t')$ is not a λ-expression or bound variable and $t|_{p.2}$ is not amongst the first n arguments of $head(t|_p)$.

Beta-reduction is defined on λ-terms as follows. A term of the form $(\lambda x.t)\,t'$ β-reduces to $t\{x \to t'\}$ where the substitution is assumed to be capture avoiding by the α-renaming of t where necessary. If term t β-reduces to t', this is symbolised $t \longrightarrow_\beta t'$.

By an overload of notation, $\|t\|$ is used to denote the length of the longest β-reduction from t. For typed λ-terms, β-reduction is terminating and confluent and a maximal strategy is known [**?**].

6.1 Beta-compatible Ordering

The ordering first compares λ-terms by the length of the longest β-reduction and then by the graceful higher-order KBO. We overload the translation $[\![\,]\!]$ to be a translation from λ-terms to untyped terms. It replaces λ-binders by a unary function symbol lam and replaces bound variables by the correct De Bruijn index. The use of De Bruijn indices ensures that $[\![]\!]$ remains an injective function. The definition of type-1 subterms requires updating as well. The translation of types is as given previously. The extended term translation follows the definition.

Definition 5 (Type-1 term). *Consider a term t of the form $(\lambda\,\overline{y_n}\,.t')\;\overline{t_n}$ or $x\;\overline{t_n}$. If there exists a position p such that $t|_p$ is a free variable, then t is a type-1 term.*

$$[\![t]\!] = \begin{cases} x & t = x \text{ and } x \text{ is free} \\ \mathsf{db}_i & t \text{ is a bound variable with } i \text{ binders} \\ & \text{between this occurrence and its binder} \\ x_t & t \text{ is a type-1 or type-2 term} \\ \mathsf{lam}\,[\![t']\!] & \text{if } t = \lambda x.t' \\ \mathsf{f}\,[\![\,\overline{\tau_n}\,]\!] & \text{if } t = \mathsf{f}\langle\,\overline{\tau_n}\,\rangle \\ [\![t_1]\!][\![t_2]\!] & \text{if } t = t_1 t_2 \end{cases}$$

Then for λ-terms, t and s, Let $t' = [\![t]\!]$ and $s' = [\![s]\!]$. We have $t >_{\beta\mathsf{kbo}} s$ if:

R1 $\|t'\| > \|s'\|$ or,
R2 $\|t'\| = \|s'\|$ and $[\![t']\!] >_{\mathsf{hb}} [\![s']\!]$

The definition of β-reduction on untyped terms is the obvious one. Most of the proofs follow almost exactly as in the combinatory case with the added complication of having to deal with terms of the form $\lambda x.t$ when doing a case analysis on terms. We therefore do not repeat them.

7 Conclusion and Discussion

We have presented an ordering that orients all ground instances of S, C, B, K and I axioms left-to-right. The ordering enjoys many other useful properties such as stability under substitution, compatibility with stable contexts, ground totality and transitivity. We hope to utilise this ordering to parameterise a complete superposition calculus for higher-order logic (HOL). For this purpose, it is conjectured that lack of full compatibility with contexts is not an obstacle. In the standard first-order proof of the completeness of superposition, compatibility with contexts is used in model construction to rule out the need for superpostion inferences beneath variables [**?**]. Thus, by utilising $>_{\mathsf{ski}}$, some superposition will most likely be required beneath variables. However, because terms with functional heads are compatible with all contexts, such inference may be quite restricted.

A complete calculus for HOL already exists, having been developed by Bentkamp et al. [**?**]. The work is an outstanding achievement, but suffers from a few drawbacks. They are forced to deal with β-equivalence classes of terms because β-reduction may actually increase the size of terms with respect to the ordering utilised. Because terms are compared modulo β-reduction, terms such as $x\,\mathsf{a}$ and $x\,\mathsf{b}$ are necessarily incomparable. Consider the substitution $\{x \to \lambda y.\mathsf{c}\}$.

The $>_{\mathsf{ski}}$ ordering presented here (or the $>_{\beta\mathsf{kbo}}$ ordering) allows weak reduction (or β-reduction) to be treated as part of the superposition calculus. This allows terms t and t' such that $t \longrightarrow_w^+ t'$ (or $t \longrightarrow_\beta^+ t'$) to be considered separate terms. Therefore, a wider range of non-ground terms become comparable. In particular the terms $t = x\,\mathsf{a}$ and $s = x\,\mathsf{b}$ can be compared. Since for any substitution θ, $\|t\theta\| = \|s\theta\|$, t and s are compared using $>_{\mathsf{hb}}$ with stability under substitution ensured by the stability under substitution of $>_{\mathsf{hb}}$.

Perhaps more importantly, a calculus based on oriented combinators would be far closer to standard first-order superposition. Terms would not contain binders and unification would be first-order. This would allow the re-use of much of the existing research with regards to data-structures and algorithms for first-order superposition provers [**?**][**?**].

Many of the definitions that have been provided here are conservative and can be tightened to allow the comparison of a far larger class of non-ground terms without losing stability under substitution. In further work, we hope to thoroughly explore these refinements.

The definition of a stable subterm can be further sharpened to allow compatibility with a greater number of contexts. We provide the intuitive idea. Consider the context $\mathsf{K}\,(\mathsf{S}\,[])\,s$. This is not a stable context because the innermost head symbol to the hole that is not a partially applied combinator is the fully applied K. However, for terms t and t' such that $\|t\| > \|t'\|$, we must have that $\|\mathsf{K}\,(\mathsf{S}\,t)\,s\| > \|\mathsf{K}\,(\mathsf{S}\,t')\,s\|$ because the terms t and t' can never become applied on a longest reduction path. This suggests the following improvement to the definition of instability.

Definition 6 (Stable Subterm). *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* ***or a*** K ***applied to two arguments or an*** I ***applied to a single argument*** *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.*

In future work, we plan to investigate improvements such as the above. However, we feel that the ordering in the form presented here is a strong base for starting work on a complete superposition calculus for HOL based on the combinatory calculus.