An AC-Compatible Knuth-Bendix Order

Konstantin Korovin[1] and Andrei Voronkov[2]

[1] MPI für Informatik
korovin@mpi-sb.mpg.de
[2] University of Manchester
voronkov@cs.man.ac.uk

Abstract. We introduce a family of AC-compatible Knuth-Bendix simplification orders which are AC-total on ground terms. Our orders preserve attractive features of the original Knuth-Bendix orders such as existence of a polynomial-time algorithm for comparing terms; computationally efficient approximations, for instance comparing weights of terms; and preference of light terms over heavy ones. This makes these orders especially suited for automated deduction where efficient algorithms on orders are desirable.

1 Introduction

Simplification orders are used in automated reasoning for pruning the search space of theorem provers and in rewriting for proving termination of rewrite rule systems and for finding complete sets of rewrite rules. E-compatible simplification orders for various equational theories E can be used for building-in equational theories in theorem provers and rewriting modulo equational theories.

Among various equational theories, theories axiomatized by the axioms of associativity and commutativity, so-called AC-theories, play a special role. Such theories very often occur in applications and require special treatment in automated systems, where AC-compatible simplification orders is a crucial ingredient. Importance of AC-compatible simplification orders triggered a huge amount of research aimed to design such orders [17–19, 16, 5, 3, 8, 4, 10, 9, 12, 11, 2, 15]. Usually, E-compatible simplification orders are designed from known simplification orders. Recently, a lot of work has been done to modify recursive path orders to obtain AC-compatible simplification orders total on ground terms [17–19, 10, 9, 12, 11]. Despite the fact that the Knuth-Bendix orders are widely used in automated deduction, to our knowledge there have been no AC-compatible simplification variant of the Knuth-Bendix order known. (There was an attempt to introduce such an order in [20] but this order is lacking the crucial monotonicity property, as we will show later).

In this paper we define a family of AC-compatible Knuth-Bendix orders $\succ_{KBO}$. These orders enjoy attractive features of the standard Knuth-Bendix orders, for example

1. a polynomial-time algorithm for term comparison;

2. computationally efficient approximations based on weight comparison, so in many practical cases we do not need to traverse the whole term each time to compare it with another term;
3. light terms are smaller than heavier ones.

Our approach share some ideas with the AC-RPO of Rubio [18, 19], but a careful exploitation of some properties of weight functions enable us to avoid complications leading to an exponential behavior in the AC-RPO case. We had to omit several proofs about $\succ_{KBO}$ on non-ground terms due to a lack of space, however, their proofs are similar to the proofs for the ground case.

2 Preliminaries. Standard Knuth-Bendix Order

In this section we introduce some standard notation and definitions.

2.1 Terms and Orders

A *signature* is a finite set of function symbols with associated arities. In this paper we assume an arbitrary but fixed signature Σ. *Constants* are function symbols of the arity 0. We assume that Σ contains at least one constant. We denote variables by x, y, z and terms by r, s, t. If a term t has the form $g(t_1, \ldots, t_n)$, where $n \geq 0$, then g is called the *top symbol* of t, denoted by $top(t)$, and $t_1, \ldots, t_n$ the *arguments* of t. We define the top symbol of a variable x to be x itself.

We use the standard notion of a *position* in a term. If π is a position in a term t and s is a term, we denote by $t[s]_\pi$ the term obtained from t by replacing its subterm at the position π by s. We will simply write $t[s]$ instead of $t[s]_\pi$ when π is fixed.

Finite multisets are defined as usual. We will only use finite multisets in this paper, so from now on a multiset always means a finite multiset. We use notation $\dot{\{}t_1, \ldots, t_n\dot{\}}$ to denote multisets. For example, $\dot{\{}a, a, b\dot{\}}$ is a multiset with two occurrences of a and one occurrence of b. The *multiset difference* of multisets A and B is denoted by $A \dot{-} B$. We write $a \mathrel{\dot{\in}} A$ to denote that a is a member of a multiset A. We use notation $\dot{\{}a \mathrel{\dot{\in}} A \mid C\dot{\}}$, to denote the submultiset of A consisting of the elements of A satisfying C.

Let $>$ be a binary relation on a set S. A *multiset extension* of $>$, denoted by $>^{mul}$, is a binary relation on multisets over S defined as follows. Let A and B be two multisets. Denote $A' = A \dot{-} B$ and $B' = B \dot{-} A$. Then $A >^{mul} B$ if A' is non-empty and for every $b \in B'$ there exists $a \in A'$ such that $a > b$. The following fact due to [6] is well-known.

LEMMA 1. *If $>$ is an order, then so is $>^{mul}$. If $>$ is a total order, then so is $>^{mul}$. If $>$ is a well-founded order, then so is $>^{mul}$.* □

Let $>$ be a binary relation on a set S. A *lexicographic extension* of $>$, denoted by $>^{lex}$, is a relation on tuples of elements of S defined as follows. Let $\bar{a} = (a_1, \ldots, a_m)$ and $\bar{b} = (b_1, \ldots, b_n)$ be two tuples. Then $\bar{a} >^{lex} \bar{b}$ if one of the following conditions holds:

1. $m > n$;
2. $m = n$ and there exists i such that $1 \leq i \leq m$, $a_i > b_i$, and for all $j \in \{1, \ldots, i-1\}$ we have $a_j = b_j$.

The following fact is not hard to check, see, e.g., [1].

LEMMA 2. *If $>$ is an order, then so is $>^{lex}$. If $>$ is a total order, then so is $>^{lex}$. If $>$ is a well-founded order, then so is $>^{lex}$.* □

A binary relation $>$ is called a *simplification order* if it is an order and it satisfies the following two properties:

1. *monotonicity*: if $s > t$, then $r[s] > r[t]$.
2. *subterm property*: if $r[s] \neq s$, then $r[s] > s$.

For every pre-order $\geq$ we denote by $>$ the corresponding strict order $>$ defined as follows: $s > t$ if and only if $s \geq t$ and $t \not\geq s$. We will use this notation for various pre-orders, for example $\succ$ will denote the strict version of $\succeq$.

Let $\geq_1, \geq_2$ be pre-orders. We call the *lexicographic product* of $\geq_1$ and $\geq_2$, denoted $\geq_1 \otimes \geq_2$, the relation $\geq$ defined as follows: $s \geq t$ if and only if either $s >_1 t$, or $s \geq_1 t$ and $s \geq_2 t$. It is not hard to argue that $\geq_1 \otimes \geq_2$ is a pre-order. We define lexicographic product $>_1 \otimes >_2$ of strict parts of $\geq_1, \geq_2$ as the strict part of $\geq_1 \otimes \geq_2$.

We will also consider lexicographic products of more than two orders.

LEMMA 3. *If $>_1, >_2$ are orders, then so is $>_1 \otimes >_2$. If $>_1, >_2$ are total orders, then so is $>_1 \otimes >_2$. If $>_1, >_2$ are well-founded orders, then so is $>_1 \otimes >_2$.* □

In our proofs below we will often compose the multiset order, the lexicographic extension, and the lexicographic product of various orders and use Lemmas 1, 2 and 3 to establish properties of the compositions.

2.2 Knuth-Bendix Order

Denote the set of natural numbers by $\mathbb{N}$. We call a *weight function* on Σ any function $w : \Sigma \to \mathbb{N}$ such that $w(a) > 0$ for every constant a. A *precedence relation* on Σ is any linear order $\gg$ on Σ. We say that a precedence relation $\gg$ is *compatible* with a weight function w if, whenever f is a unary function symbol and $w(f) = 0$, then f is the greatest element of Σ w.r.t. $\gg$.

The definition of the Knuth-Bendix order on the set of ground terms of the signature Σ is parameterized by (i) a weight function w on Σ; and (ii) a precedence relation $\gg$ on Σ compatible with w. The compatibility condition ensures that the Knuth-Bendix order is a simplification order total on ground terms, see, e.g., [1]. In this paper, f will always denote a unary function symbol of the weight 0.

In the sequel we assume a fixed weight function w on Σ and a fixed precedence relation $\gg$ on Σ. We call $w(g)$ the *weight* of g. The *weight* of any ground term

t, denoted $|t|$, is defined as follows: for any constant c we have $|c| = w(c)$ and for any function symbol g of a positive arity $|g(t_1, \ldots, t_n)| = w(g) + |t_1| + \ldots + |t_n|$.

The *Knuth-Bendix order induced by* w and $\gg$ is the binary relation $\succ_{KBO}$ on ground terms defined as follows. For any ground terms $t = g(t_1, \ldots, t_n)$ and $s = h(s_1, \ldots, s_k)$ we have $t \succ_{KBO} s$ if one of the following conditions holds:

1. $|t| > |s|$;
2. $|t| = |s|$ and $g \gg h$;
3. $|t| = |s|$, $g = h$ and $(t_1, \ldots, t_n) \succ_{KBO}^{lex} (s_1, \ldots, s_n)$.

It is known that for every weight function w and precedence relation $\gg$ compatible with w, the Knuth-Bendix order induced by w and $\gg$ is a simplification order total on ground terms (see e. g. [1]).

2.3 AC-compatible orders

Let E be an equational theory and $>$ be a partial order on ground terms of a signature Σ. Denote equality with respect to E by $=_E$. We say that an order $>$ is *E-compatible* if it satisfies the following property: if $s > t$, $s =_E s'$ and $t =_E t'$, then $s' > t'$. The order $>$ is called E-total, if for all ground terms s, t, if $s \neq_E t$, then either $s > t$ or $t > s$.

Let $+$ be a binary function symbol. The *AC-theory* for $+$ is the equational theory axiomatized by set of two formulas

$$\forall x \forall y \forall z((x + y) + z \simeq x + (y + z));$$
$$\forall x \forall y(x + y \simeq y + x).$$

From now on we assume that we are given a fixed signature Σ with a distinguished subset Σ_{AC} of binary function symbols. The members of Σ_{AC} will be called *AC-symbols*. Two terms s, t are called *AC-equal*, denoted $s =_{AC} t$, if they are equal in the equational theory generated by the union of the AC-theories for all $g \in \Sigma_{AC}$. An order is called *AC-compatible* if it is E-compatible with respect to this equational theory.

2.4 Main results

Our main aim is to find an AC-compatible AC-total simplification order which generalizes the standard Knuth-Bendix order for the case of AC-theories. In the rest of this paper we define a family of such orders, each order $\succ_{KBO}$ in this family is induced by a weight function w and a precedence relation $\gg$ compatible with w. We prove the following results.

1. $\succ_{KBO}$ is an AC-compatible AC-total simplification order,
2. On the terms without AC-symbols, $\succ_{KBO}$ coincides with the standard Knuth-Bendix order induced by w and $\gg$.
3. If Σ contains no unary function symbols of the weight 0, then for every ground term t there exists a finite number of terms s such that $t \succ_{KBO} s$.

Further, we extend the orders $\succ_{KBO}$ to non-ground terms in such a way that for all terms s, t and substitutions θ, if $s \succ_{KBO} t$, then $s\theta \succ_{KBO} t\theta$.

3 The Ground Case

3.1 Flattened terms

In the sequel the symbol $+$ will range over Σ_{AC}. Let us call a term *normalized* if it has no subterms of the form $(r+s)+t$. Evidently, every term is AC-equal to a normalized term. Since we aim at finding AC-compatible simplification orders, it is enough for us to define these orders only for normalized terms. For normalized terms, we introduce a special well-known notation, called *flattened term*.

To this end, we consider all AC-symbols to be varyadic, i.e., having an unbounded arity greater than or equal to 2. A term s using the varyadic symbols is called *flattened* if for every non-variable subterm t of s, if t has the form $+(t_1,\ldots,t_n)$, then the top symbols of $t_1,\ldots,t_n$ are distinct from $+$. We identify a subterm $+(t_1,\ldots,t_m)$ with the normalized term $(t_1+(t_2+\ldots+t_n))$. We will sometime write subterms of flattened terms as $t_1+\ldots+t_n$. In the sequel we will only deal with flattened terms.

Note that we have to be careful with defining substitutions into flattened terms and the subterm property for them. When we substitute a term $s_1+\ldots+s_m$ for a variable x in $x+t_1+\ldots+t_n$, we obtain $s_1+\ldots+s_m+t_1+\ldots+t_n$. To prove the subterm property for an order $>$ on ordinary terms, we also have to prove the following *cancellation property* for flattened terms: $s_1+s_2+\ldots+s_n > s_2+\ldots+s_n$.

Similarly, we have to be careful with defining weights of terms with varyadic symbols. We want the weight to be invariant under $=_{AC}$, in particular, the weight of a term must coincide with the weight of a flattened term equal to it modulo AC. Therefore, we modify the definition of weight as follows.

Definition 4. (Weight) The *weight* of a ground term t, denoted $|t|$, is defined as follows. Let $t=g(t_1,\ldots,t_n)$, where $n\geq 0$. Then

1. if $g\notin\Sigma_{AC}$, then $|t|=w(g)+|t_1|+\ldots+|t_n|$.
2. if $g\in\Sigma_{AC}$, then $|t|=(n-1)w(g)+|t_1|+\ldots+|t_n|$. □

We have the following straightforward result.

Lemma 5. *Let r,s,t be terms. If $|s|=|t|$, then $|r[s]|=|r[t]|$. Likewise, if $|s|>|t|$, then $|r[s]|>|r[t]|$.* □

3.2 Relation $\succ_+$

All relations introduced below will be AC-compatible. Therefore, in the sequel we will consider the AC-equality instead of the syntactic equality and consider relations on the equivalence classes modulo $=_{AC}$.

To define an AC-compatible weight-based simplification order, let us first define, for each AC-symbol $+$, an auxiliary partial order $\succ_+$ on multisets of flattened terms.

First we introduce the following *pre-order* $\geq_{top}$ on terms: $s\geq_{top} t$ if and only if $top(s)\gg top(t)$ or $top(s)=top(t)$. Note that this order is also defined for non-ground terms. Likewise, we introduce the pre-order $\geq_w$ on ground terms as follows: $s \geq_w t$ if $|s| \geq |t|$. Naturally, the strict versions of $\geq_{top}$ and $\geq_w$ are denoted by $>_{top}$ and $>_w$, respectively.

Definition 6. (Relation $\succ_+$) Let M, N be two multisets of flattened ground terms and let

$$M' = \dot{\{}t \mathrel{\dot\in} M \mid top(t) \gg +\dot{\}};$$
$$N' = \dot{\{}t \mathrel{\dot\in} N \mid top(t) \gg +\dot{\}}.$$

We define $M \succeq_+ N$ if and only if

$$M' \; (\geq_w \otimes \geq_{top})^{mul} \; N'.$$

□

In other words, we can define the order $\succ_+$ as follows. First, remove from M and N all elements with top symbols smaller than or equal to $+$. Then compare the remaining multisets using the multiset order in which the terms are first compared by weight and then by their top symbol.

Lemma 7. *For each symbol $+ \in \Sigma_{AC}$ the relation $\succ_+$ is a well-founded order.*

Proof. Follows immediately from the observation that the strict part of $(\geq_w \otimes \geq_{top})^{mul}$ is a well-founded order (by Lemmas 1 and 3). □

Let us give a characterization of the relation $\succ_+$. Let M be a multiset of ground terms and v be a positive integer. Denote by $selected(+, v, M)$ the multiset of top functors of all terms in M of the weight v whose top symbol is greater than $+$ w.r.t. $\gg$. Then we have $M \succ_+ N$ if and only if there exists an integer v such that $selected(+, v, M) >^{mul}_{top} selected(+, v, N)$ and for all $v' > v$, $selected(+, v', M) =^{mul} selected(+, v', N)$. Let $\equiv_+$ denote the incomparability relation on multisets of terms w.r.t. $\succ_+$. That is, given two multisets M, N, we have $M \equiv_+ N$ if and only if neither $M \succ_+ N$ nor $N \succ_+ M$. Now it is easy to check that two multisets of terms M and N are incomparable w.r.t. $\succ_+$ if and only if for each weight v we have $selected(+, v, M) = selected(+, v, N)$ and therefore $\equiv_+$ is indeed an equivalence relation on terms. So $\succ_+$ can be seen as a total well-founded order on the equivalence classes of multisets modulo $\equiv_+$.

3.3 Order $\succ_{KBO}$

Using the relation $\succ_+$, we can define an AC-compatible simplification order $\succ_{KBO}$.

Definition 8. (Order $\succ_{KBO}$) Let $t = h(t_1, \ldots, t_n)$ and $s = g(s_1, \ldots, s_k)$ be flattened ground terms. Then $t \succ_{KBO} s$ if and only if one of the following conditions holds:

1. $|t| > |s|$; or

2. $|t| = |s|$ and $h \gg g$; or
3. $|t| = |s|$, $h = g$, and either
   (a) $h \notin \Sigma_{AC}$ and $(t_1, \dots, t_n) \succ_{KBO}^{lex} (s_1, \dots, s_n)$; or
   (b) $h \in \Sigma_{AC}$ and
      i. $\dot{\{}t_1, \dots, t_n\dot{\}} \succ_h \dot{\{}s_1, \dots, s_k\dot{\}}$; or
      ii. $\dot{\{}t_1, \dots, t_n\dot{\}} \equiv_h \dot{\{}s_1, \dots, s_k\dot{\}}$ and $n > k$; or
      iii. $\dot{\{}t_1, \dots, t_n\dot{\}} \equiv_h \dot{\{}s_1, \dots, s_k\dot{\}}$, $n = k$ and $\dot{\{}t_1, \dots, t_n\dot{\}} \succ_{KBO}^{mul} \dot{\{}s_1, \dots, s_k\dot{\}}$. □

Let us remark that similar to the AC-RPO of Rubio [18, 19] we make a special treatment of the immediate subterms below $+$ having top symbols greater than $+$. To this end, we use the relation $\succ_+$, which allows us to avoid recursive computations deeper into subterms at this stage (we need only to compare weights and top symbols of the immediate subterms). As a result, we gain some efficiency. More importantly, using properties of the weight functions we can avoid the exponential behavior of AC-RPO caused by enumerating embeddings of certain subterms.

Lemma 9. *$\succ_{KBO}$ is an AC-compatible AC-total order on ground terms.*

Proof. It is easy to see that $\succ_{KBO}$ is AC-compatible. The AC-totality can be proved by a routine induction on terms.

Let us prove that $\succ_{KBO}$ is an order. Let us call the *f-height* of a term r, denoted by $height_f(r)$, the greatest number n such that $r = f^n(r')$. The proof is by induction on the order $>'$ on ground terms defined as follows: $t >' s$ if $|t| > |s|$ or $|t| = |s|$ and $height_f(t) > height_f(s)$. Obviously, $>'$ is the lexicographic product of two well-founded orders, and so a well-founded order itself.

Note the following property of $>'$: if $t >' s$, then $t \succ_{KBO} s$. Therefore, it is enough to prove that for each pair of natural numbers (k, l), the relation $\succ_{KBO}$ is an order on the set of ground terms

$$\{t \mid |t| = k \text{ and } height_f(t) = l\}.$$

But this follows from the following observation: $\succ_{KBO}$ on this set of terms is defined as a lexicographic product of the following five orders:

$$
\begin{aligned}
&t >_1 s \Leftrightarrow h \gg g;\\
&t >_2 s \Leftrightarrow (t_1, \dots, t_n) \succ_{KBO}^{lex} (s_1, \dots, s_n) \text{ and } h = g \notin \Sigma_{AC};\\
&t >_3 s \Leftrightarrow \dot{\{}t_1, \dots, t_n\dot{\}} \succ_h \dot{\{}s_1, \dots, s_k\dot{\}} \text{ and } h = g \in \Sigma_{AC};\\
&t >_4 s \Leftrightarrow n > k \text{ and } h = g \in \Sigma_{AC};\\
&t >_5 s \Leftrightarrow \dot{\{}t_1, \dots, t_n\dot{\}} \succ_{KBO}^{mul} \dot{\{}s_1, \dots, s_k\dot{\}} \text{ and } h = g \in \Sigma_{AC}.
\end{aligned}
$$

Note that $\succ_{KBO}^{lex}$ and $\succ_{KBO}^{mul}$ used in this definition are orders by the induction hypothesis and by Lemmas 2 and 1. □

THEOREM 10. *The relation $\succ_{KBO}$ is an AC-compatible AC-total simplification order on ground terms.*

PROOF. By Lemma 9, $\succ_{KBO}$ is an order, so it only remains to prove the subterm property, cancellation property, and monotonicity. The cancellation property is obvious, since $|s_0 + s_1 + \ldots + s_n| > |s_1 + \ldots + s_n|$. The subterm property is checked in the same way as for the standard Knuth-Bendix order.

Let us prove the monotonicity. By Lemma 9, $\succ_{KBO}$ is an AC-compatible AC-total order. In particular, $\succ_{KBO}$ is transitive, so it remains to prove the following property: if $t \succ_{KBO} s$, then for every function symbol g we have $g(r_1, \ldots, r_{i-1}, t, r_{i+1}, \ldots, r_n) \succ g(r_1, \ldots, r_{i-1}, s, r_{i+1}, \ldots, r_n)$. When $g \notin \Sigma_{AC}$, the proof is identical to that for the standard Knuth-Bendix order, so we only consider the case when g is an AC-symbol $+$.

We have to prove the following statement for all terms $s, t, r_1, \ldots, r_m$: let $u = t + r_1 + \ldots + r_m$ and $v = s + r_1 + \ldots + r_m$, then $t \succ_{KBO} s$ implies $u \succ_{KBO} v$. Let $t = h(t_1, \ldots, t_n)$ and $s = g(s_1, \ldots, s_k)$. Consider all possible cases of Definition 8 of $\succ_{KBO}$.

1. $|t| > |s|$. In this case by Lemma 5 we have $|u| > |v|$, and so $u \succ_{KBO} v$.
   Now we can assume $|t| = |s|$, hence by Lemma 5 $|u| = |v|$. Denote by U and V the multisets of arguments of u and v, respectively. Note that U is not necessarily equal to $\dot{\{}t, r_1, \ldots, r_m\dot{\}}$: indeed, the top symbol of t may be $+$, and then we have to flatten $t + r_1 + \ldots + r_m$ to obtain the arguments of u. Likewise, V is not necessarily equal to $\dot{\{}s, r_1, \ldots, r_m\dot{\}}$. Denote by p, q the number of elements in U, V respectively. Note that
   $$p = \begin{cases} m+1, \text{ if } top(t) \neq +; \\ m+n, \text{ if } top(t) = +. \end{cases}$$
   $$q = \begin{cases} m+1, \text{ if } top(s) \neq +; \\ m+k, \text{ if } top(s) = +. \end{cases}$$
   Since $|u| = |v|$ and $top(u) = top(v) = +$, the comparison of u and v should be done using clauses (3(b)i)–(3(b)iii) of Definition 8. That is, first we check $U \succ_+ V$. Then, if $U \equiv_+ V$, we check if $p > q$. Finally, if $p = q$, we compare U and V using the multiset order $\succ_{KBO}^{mul}$. Consider the remaining cases.
2. $h \gg g$. Let us show that if $h \gg +$ then $U \succ_+ V$ and so $u \succ_{KBO} v$. If $+ \gg g$ then we have $U \succ_+ U \dot{-} \dot{\{}t\dot{\}} = \dot{\{}r_1, \ldots, r_m\dot{\}} = V \dot{-} \dot{\{}s\dot{\}} \equiv_+ V$. If $g \gg +$ then $\dot{\{}t\dot{\}} \succ_+ \dot{\{}s\dot{\}}$ and hence $U = \dot{\{}t, r_1, \ldots, r_m\dot{\}} \succ_+ \dot{\{}s, r_1, \ldots, r_m\dot{\}} = V$. If $g = +$ then s is of the form $s_1 + \ldots + s_k$. We have $\dot{\{}t\dot{\}} \succ_+ \dot{\{}s_1, \ldots, s_k\dot{\}}$, since the weight of each arguments of s is strictly less than the weight of t, and therefore $U \succ_+ V$.
   Now if $+ \gg h$, then $U \equiv_+ V$ and $p = q$. In this case $u \succ_{KBO} v \Leftrightarrow U \succ_{KBO}^{mul} V \Leftrightarrow t \succ_{KBO} s$, so $u \succ_{KBO} v$. It remains to consider the case $h = +$. In this case we have $U \succeq_+ V \dot{-} \dot{\{}s\dot{\}} \equiv_+ V$ and either $U \succ_+ V$, so $u \succ_{KBO} v$, or we have $U \equiv_+ V$ and $p > q$, so $u \succ_{KBO} v$, by (3(b)ii) of Definition 8.
3. $h = g$.

(a) $h \neq +$. Then $U \equiv_+ V$ and $p = q$. In this case $u \succ_{KBO} v \Leftrightarrow U \succ_{KBO}^{mul} V \Leftrightarrow t \succ_{KBO} s$.

(b) Now it remains to consider the case $h = g = +$. In this case $U = \dot{\{}t_1, \ldots, t_n, r_1, \ldots, r_m\dot{\}}$ and $V = \dot{\{}s_1, \ldots, s_k, r_1, \ldots, r_m\dot{\}}$. Since $t \succ_{KBO} s$, it is enough to consider the following cases.

   i. $\dot{\{}t_1, \ldots, t_n\dot{\}} \succ_+ \dot{\{}s_1, \ldots, s_k\dot{\}}$. In this case $U \succ_+ V$, hence $u \succ_{KBO} v$.

   ii. $\dot{\{}t_1, \ldots, t_n\dot{\}} \equiv_+ \dot{\{}s_1, \ldots, s_k\dot{\}}$ and $n > k$. In this case $U \equiv_+ V$ but $p > q$, hence $u \succ_{KBO} v$.

   iii. $\dot{\{}t_1, \ldots, t_n\dot{\}} \equiv_+ \dot{\{}s_1, \ldots, s_k\dot{\}}$, $n = k$, and $\dot{\{}t_1, \ldots, t_n\dot{\}} \succ_{KBO}^{mul} \dot{\{}s_1, \ldots, s_k\dot{\}}$. In this case $U \equiv_+ V$, $p = q$, but $U \succ_{KBO}^{mul} V$, hence $u \succ_{KBO} v$.

The proof is complete. □

Suppose that Σ does not contains a unary function symbol f of the weight 0. In this case for each weight v there is only a finite number of ground terms of the weight v. Therefore, we have the following result.

PROPOSITION 11. *If Σ does not contain a unary function symbol f of the weight 0, then for every term t, there exists only a finite number of terms s such that $t \succ_{KBO} s$.* □

Now let us show that if our signature contains only two AC-symbols and in addition one of them is maximal and another is minimal w.r.t. $\gg$, then we can considerably simplify definition of AC-KBO by avoiding $\succ_h$ comparisons. In particular the following definition will satisfy all required properties.

DEFINITION 12. (Simplified AC-KBO for two AC symbols) Consider a signature Σ containing only two AC-symbols, such that one of them is maximal and another is minimal w.r.t. $\gg$ in Σ.

Let $t = h(t_1, \ldots, t_n)$ and $s = g(s_1, \ldots, s_k)$ be flattened ground terms. Then $t \succ'_{KBO} s$ if and only if one of the following conditions holds:

1. $|t| > |s|$; or
2. $|t| = |s|$ and $h \gg g$; or
3. $|t| = |s|$, $h = g$, and either
   (a) $h \notin \Sigma_{AC}$ and $(t_1, \ldots, t_n) \succ'^{\,lex}_{KBO} (s_1, \ldots, s_n)$; or
   (b) $h \in \Sigma_{AC}$ and
      i. $n > k$ and h is maximal in Σ w.r.t. $\gg$; or
      ii. $k > n$ and h is minimal in Σ w.r.t. $\gg$; or
      iii. $k = n$ and $\dot{\{}t_1, \ldots, t_n\dot{\}} \succ'^{\,mul}_{KBO} \dot{\{}s_1, \ldots, s_k\dot{\}}$. □

THEOREM 13. *The relation $\succ'_{KBO}$ is an AC-compatible AC-total simplification order on ground terms.*

PROOF. We skip the proof which is similar to the general case. □

4 Non-Ground Order

In this section we will generalize the AC-compatible Knuth-Bendix order $\succ_{KBO}$ to non-ground terms. The definition will be very similar to the ground case. We will have to change the definitions of the weight and slightly change the definition of $\succ_+$. As before, we will be dealing with flattened terms.

Let us modify the notion of weight to non-ground terms. In fact, we will introduce two different weights $|t|$ and $||t||$. As before, we assume that we are given a weight function w and a precedence relation $\gg$ compatible with w. Let e denote the constant in Σ having the least weight among all constants in Σ. It is not hard to argue that $|e|$ is also the least weight of a ground term.

DEFINITION 14. (Weight $|t|$) The *weight* of a term t, denoted $|t|$, is defined as follows.

1. If t is a variable, then $|t| = w(e)$.
2. If $t = g(t_1, \ldots, t_n)$ and $g \notin \Sigma_{AC}$, then $|t| = w(g) + |t_1| + \ldots + |t_n|$.
3. If $t = g(t_1, \ldots, t_n)$ and $g \in \Sigma_{AC}$, then $|t| = (n-1)w(g) + |t_1| + \ldots + |t_n|$. □

It is not hard to argue that the weight of a term t is equal to the weight of the ground term obtained from t by replacing all variables by e. Therefore, Lemma 5 also holds for non-ground terms.

LEMMA 15. *Let r, s, t be terms. If $|s| = |t|$, then $|r[s]| = |r[t]|$. Likewise, if $|s| > |t|$, then $|r[s]| > |r[t]|$.* □

Let t be a term. Denote by $vars(t)$ the multiset of variables of t. For example, $vars(g(x, a, h(y, x))) = \dot{\{}x, y, x\dot{\}}$.

DEFINITION 16. (Generalized Weight) A *generalized weight* is a pair (n, V), where n is a positive integer and V is a multiset of variables. Let us introduce a pre-order $\geq$ and an order $>$ on generalized weights as follows. We let $(m, M) \geq (n, N)$ if $m \geq n$ and N is a submultiset of M. We let $(m, M) > (n, N)$ if $m > n$ and N is a submultiset of M. The *generalized weight of a term t*, denoted $||t||$, is the pair $(|t|, vars(t))$. We write $t \geq_W s$ if $||t|| \geq ||s||$ and $t >_W s$ if $||t|| > ||s||$.

Note that $>_W$ is *not* a strict version of $\geq_W$. However, it is easy to see that $>_W$ is a well-founded order. The following properties of $\geq_W$ and $>_W$ are also not difficult to check.

LEMMA 17. *Let r, s, t be terms. If $s \geq_W t$, then $r[s] \geq_W r[t]$. Likewise, if $s >_W t$, then $r[s] >_W r[t]$. Moreover, if s, t are ground terms, then $s \geq_w t$ if and only if $s \geq_W t$, and $s >_w t$ if and only if $s >_W t$.* □

Note that $\geq_W$ is not a total pre-order. For example, if x, y are two different variables, then neither $x \geq_W y$ nor $y \geq_W x$ holds.

4.1 Relation $\succ_+$

Let us now generalize the relation $\succ_+$ to non-ground terms. The definition is more complex that in the ground case because of one technical problem: the order $>_W$ is not the strict version of $\geq_W$. Therefore, we cannot compose orders using $\geq_W$ to obtain new orders as before. In particular, the definition of a multiset extension of an order does not work any more and should be replaced.

First, instead of the pre-order $\geq_w \otimes \geq_{top}$ used in the definition of $\succ_+$ on ground terms, we introduce a pre-order $\geq_{W,top}$ defined as $\geq_W \otimes \geq_{top}$. We also write $s =_{W,top} t$ if $||s|| = ||t||$ and $top(s) = top(t)$. Then let us define an order $>_{W,top}$ as follows: $s >_{W,top} t$ if either $s >_W t$ or $s \geq_W t$ and $top(s) \gg top(t)$.

Now, to define an analogue of $(\geq_w \otimes \geq_{top})^{mul}$ used in the definition of $\succ_+$ for ground terms, let us define the following *deletion operation* on pairs of multisets M, N: if $t \dot{\in} M$, $s \dot{\in} N$, and $t =_{W,top} s$, then delete one occurrence of t from M and one occurrence of s from N.

Definition 18. (Relation $\succ_+$) Let M, N be two multisets of flattened terms and let

$$M' = \dot{\{} t \dot{\in} M \mid t \text{ is a variable or } top(t) \gg + \dot{\}};$$
$$N' = \dot{\{} t \dot{\in} N \mid t \text{ is a variable or } top(t) \gg + \dot{\}}.$$

Let M'', N'' be obtained by applying the deletion operation to M', N' while possible. Then we define $M \succ_+ N$ if M'' contains a non-variable term and for every $s \in N''$ there exists $t \in M''$ such that $t >_{W,top} s$. We also define $M \succeq_+ N$ if either $M \succ_+ N$ or N'' is empty and M'' contains only variables. □

Similarly to the ground case, we have the following lemma.

Lemma 19. *For each symbol $+ \in \Sigma_{AC}$ the relation $\succ_+$ is a well-founded order. Moreover, on ground terms it coincides with the order $\succ_+$ of Definition 6.* □

4.2 Order $\succ_{KBO}$

Using the relation $\succ_+$, we can define an AC-compatible simplification order $\succ_{KBO}$ in essentially the same way as for ground terms.

Definition 20. (Order $\succ_{KBO}$) Let us define the relation $\succ_{KBO}$ for non-ground terms as follows. If x is a variable, then for every term s it is not true that $x \succ_{KBO} s$. If y is a variable then $t \succ_{KBO} y$ if and only if y occurs in t and is distinct from t. Let $t = h(t_1, \ldots, t_n)$ and $s = g(s_1, \ldots, s_k)$ be flattened terms. Then $t \succ_{KBO} s$ if and only if one of the following conditions holds:

1. $t >_W s$; or
2. $t \geq_W s$ and $h \gg g$; or
3. $t \geq_W s$, $h = g$, and either
   (a) $h \notin \Sigma_{AC}$ and $(t_1, \ldots, t_n) \succ_{KBO}^{lex} (s_1, \ldots, s_n)$; or

(b) $h \in \Sigma_{AC}$ and
   i. $\dot{\{}t_1, \ldots, t_n\dot{\}} \succ_h \dot{\{}s_1, \ldots, s_k\dot{\}}$; or
   ii. $\dot{\{}t_1, \ldots, t_n\dot{\}} \succeq_h \dot{\{}s_1, \ldots, s_k\dot{\}}$ and $n > k$; or
   iii. $\dot{\{}t_1, \ldots, t_n\dot{\}} \succeq_h \dot{\{}s_1, \ldots, s_k\dot{\}}$, $n = k$ and $\dot{\{}t_1, \ldots, t_n\dot{\}} \succ_{KBO}^{mul} \dot{\{}s_1, \ldots, s_k\dot{\}}$. □

THEOREM 21. *The relation $\succ_{KBO}$ is an AC-compatible simplification order. Moreover, on ground terms it coincides with the order of Definition 8.* □

THEOREM 22. *$\succ_{KBO}$ is closed under substitutions, that is, if $t \succ_{KBO} s$, then for every substitution θ, $t\theta \succ_{KBO} s\theta$.* □

5 Related Work

In general, the Knuth-Bendix order and recursive path orders are incomparable in the sense that there are rewrite (equational) systems that can be oriented by an instance of the Knuth-Bendix order but cannot be oriented by recursive path orders, and vice versa. To compare the Knuth-Bendix order with orders based on polynomial interpretations (or combinations of polynomial interpretations with recursive path orders) let us note that usually it is difficult to find a suitable polynomial interpretation which orients a given rewrite (equational) system. For the Knuth-Bendix order, we can employ some known efficient algorithms [7, 13, 14].

An attempt to define an AC-compatible Knuth-Bendix order was undertaken in [20] for a special case when each AC-symbol $+$ is of the weight 0 and is also a maximal symbol w.r.t. $\gg$. It is proposed to compare terms with the top symbol $+$ first by weight and then by comparing the multisets of their arguments. Let us give an example demonstrating that the order defined in this way lacks the monotonicity property.

Consider the weight function w such that $w(+) = 0$ and $w(c) = w(d) = w(g) = 1$ and a precedence relation $\gg$ such that $+ \gg g$. Let $t = c + d$ and $s = g(c)$. Then $|t| = |s|$, and therefore $t \succ_{KBO} s$. Take any term r. Then by monotonicity we must have $r + c + d \succ_{KBO} r + g(c)$. But in fact we have $r + g(c) \succ_{KBO} r + c + d$, since $|g(c)| > |c|$ and $|g(c)| > |d|$.

References

1. F. Baader and T. Nipkow. *Term Rewriting and All That*. Cambridge University press, Cambridge, 1998.
2. L. Bachmair. Associative-commutative reduction orderings. *Information Processing Letters*, 43(1):21–27, 1992.
3. L. Bachmair and D. Plaisted. Associative path orderings. In *Rewriting Techniques and Applications*, volume 202 of *Lecture Notes in Computer Science*, pages 241–254, 1985.

4. A. B. Cherifa and P. Lescanne. Termination of rewriting systems by polynomial interpretations and its implementation. *Science of Computer Programming*, 9(2):137–159, 1987.
5. N. Dershowitz, J. Hsiang, A. Josephson, and D. Plaisted. Associative-commutative rewriting. In *Proc. International Joint Conference on Artificial Intelligence (IJCAI)*, pages 940–944, 1983.
6. N. Dershowitz and Z. Manna. Proving termination with multiset orderings. *Communications of the ACM (CACM)*, 22(8):465–476, 1979.
7. J. Dick, J. Kalmus, and U. Martin. Automating the Knuth-Bendix ordering. *Acta Informatica*, 28(2):95–119, 1990.
8. I. Gnaedig and P. Lescanne. Proving termination of associative commutative rewriting systems by rewriting. In *International Conference on Automated Deduction*, volume 230 of *Lecture Notes in Computer Science*, pages 52–61, 1986.
9. D. Kapur and G. Sivakumar. A total, ground path ordering for proving termination of ac-rewrite systems. In *Rewriting Techniques and Applications*, volume 1232 of *Lecture Notes in Computer Science*, pages 142–156, 1997.
10. D. Kapur and G. Sivakumar. Proving associative-communicative termination using rpo-compatible orderings. In *Automated Deduction in Classical and Non-Classical Logics, Selected Papers*, volume 1761 of *Lecture Notes in Computer Science*, pages 39–61, 1998.
11. D. Kapur, G. Sivakumar, and H. Zhang. A new method for proving termination of ac-rewrite systems. In *Foundations of Software Technology and Theoretical Computer Science (FSTTCS)*, volume 472 of *Lecture Notes in Computer Science*, pages 133–148, 1990.
12. D. Kapur, G. Sivakumar, and H. Zhang. A path ordering for proving termination of ac rewrite systems. *Journal of Automated Reasoning*, 14(2):293–316, 1995.
13. K. Korovin and A. Voronkov. Verifying orientability of rewrite rules using the Knuth-Bendix order. In *In Proc. 10th International Conference on Rewriting Techniques and Applications (RTA)*, volume 2051 of *Lecture Notes in Computer Science*, pages 137–153. Springer, 2001. full version to appear in Information and Computation.
14. K. Korovin and A. Voronkov. Orienting equalities using the Knuth–Bendix order. In *Proceedings of the 18th IEEE Symposium on Logic in Computer Science (LICS'03)* to appear, 2003.
15. Claude Marché. Associative-commutative reduction orderings via head-preserving interpretations. Technical Report 95-2, E.N.S. de Cachan, 1995.
16. P. Narendran and M. Rusinowitch. Any ground associative-commutative theory has a finite canonical system. In *Rewriting Techniques and Applications*, volume 488 of *Lecture Notes in Computer Science*, pages 423–434, 1991.
17. R. Nieuwenhuis and A. Rubio. A precedence-based total AC-compatible ordering. In *Rewriting Techniques and Applications*, volume 690 of *Lecture Notes in Computer Science*, pages 374–388, 1993.
18. A. Rubio. A fully syntactic AC-RPO. to appear in Information and Computation.
19. A. Rubio. A fully syntactic AC-RPO. In *Rewriting Techniques and Applications*, volume 1631 of *Lecture Notes in Computer Science*, pages 133–147, 1999.
20. J. Steinbach. AC-termination of rewrite systems: A modified Knuth-Bendix ordering. In H. Kirchner and W. Wechler, editors, *Algebraic and Logic Programming, Second International Conference, Nancy, France, October 1-3, 1990, Proceedings*, volume 463 of *Lecture Notes in Computer Science*. Springer Verlag, 1990.