Extension orderings revisited*

Jean-Pierre Jouannaud

LIX, École Polytechnique
91400 Palaiseau, FRANCE
Email: jouannaud@lix.polytechnique.fr
http://www.lix.polytechnique.fr/Labo/jouannaud

Abstract. The goal of this paper is threefold : to revisit the definitions of order-preserving functionals, to make precise whether well-founded quasi-orderings and well quasi-orderings are closed under these functionals, and to provide with proofs which are simple enough to be taught to undergraduate students.
It appears from the recent literature that some of these notions are misunderstood, and that the most recent simple proof methods of some important results are not known as they should, even from the authors of textbooks. We hope that this paper will remedy to this situation.

1 Introduction

Well-founded orderings are one of the everyday tool of the theoretical computer scientist, when proving that some computation terminates, when showing some property by induction, or when using an automated theorem prover. Therefore, a good practice and a good understanding of well-founded orderings often makes the difference.

Building well-founded orderings is a difficult task that often requires a *divide and conquer* approach. There are two families of techniques. The first, combining two well-founded orderings into their union, is not covered here although extremely useful and simple. We concentrate on the second, aiming at extending a well-founded ordering on a set into well-founded orderings on data structures built from elements of that set. More precisely, we review the standard constructions of well-founded orderings on tuples, words, multisets and trees. We emphasize:

1. simple, easy to understand, carefull definitions;
2. quasi-orderings, which are more important in practice than orderings;
3. well-founded quasi-orderings versus well quasi-orderings;
4. and the simplicity of proofs.

* This work was partly supported by the RNTL project AVERROES

It appears that well-founded quasi-orderings are more flexible, easier to construct in pratice, fit well with lexicographic compositions. Moreover, proving them well-founded is by no means harder than proving orderings to be well-founded.

Besides, we will show that all difficult well-orderedness properties known as Dicson's lemma, Higman's lemma and Kruskal's theorem are useless when it comes to well-foundedness proofs, and may even require stronger hypotheses than needed in this respect.

Finally, we give some new material regarding Dershowitz's recursive path (quasi-) ordering [4]. This new material comes in the form of more elegant and simpler proofs than in the literature, although some of them appeared in the more general context of higher-order calculi [7].

Missing notations, definitions and commutation results introducing to other termination methods can be found in [5].

2 Quasi-orderings

The goal of this section is to extend to arbitrary non-transitive relations notions which are usually given for quasi-orderings. We start with general definitions about relations.

Definition 1. *A binary relation $\succeq$ on a set S is*

- total *if any two elements of S are comparable ;*
- reflexive *if $\forall x \in S\ x \succeq x$;*
- antireflexive *if $\nexists x \in S\ x \succeq x$;*
- symmetric *if $\forall x, y \in S\ x \succeq y \Rightarrow y \succeq x$;*
- anti-symetric *if $x = y\ \forall x, y \in S$ such that $x \succeq y$ and $y \succeq x$;*
- transitive *if $\forall x, y, z \in S$ such that $x \succeq y$ and $y \succeq z$, then $x \succeq z$;*
- *an* equivalence *if it is reflexive, symmetric and transitive ;*
- *a* quasi-ordering *if it is reflexive and transitive ;*
- *an* ordering *if it is reflexive, transitive and anti-symetric ;*
- *a* strict ordering *if it is anti-reflexive and transitive ;*

Definition 2. *Given an arbitrary binary relation $\succeq$ on a set S, we define:*

- *its* inverse *$\preceq$ such that $a \preceq b$ iff $b \succeq a$;*
- *its* equivalence *$\simeq$ such that $a \simeq b$ iff $a \succeq b$ and $b \succeq a$;*
- *its* strict part *$\succ$ such that $a \succ b$ iff $a \succeq b$ and $b \not\succeq a$.*
- *its reflexive transitive closure $\succeq^*$*
- *its transitive closure $\succeq^+$*

The strict part of a relation coincides with the relation itself in absence of loops. It therefore coincides with the usual notion when the relation is transitive. So does the equivalence associated to a transitive relation:

Property 1. Let $\succeq$ be a quasi-ordering of S. Then its strict part $\succ$ is a strict ordering and its equivalence $\simeq$ is an equivalence.

Definition 3. *Given a binary relation $\succ$ on a set S, an element $x \in S$ is*

- *in* normal form *if there exists no element $y \in S$ such that $x \succ y$;*
- normalizing *if there exists a finite sequence of elements of S issuing from x and ending in a normal form y, that is $x = x_0 \succ \ldots \succ x_n = y$;*
- strongly normalizing *if there exists no infinite sequence of elements of S issuing from x, that is $x = x_1 \succ x_2 \succ \ldots \succ x_n \succ \ldots$;*

Definition 4. *A relation $\succeq$ on S is well-founded if any element of S is strongly normalizing for its strict part $\succ$.*

This definition is usually given for quasi-orderings. We will use it for arbitrary relations. The difference, however, is small :

Property 2. Let $\succeq$ be a well-founded relation. Then, its transitive closure $\succeq^*$ is a well-founded quasi-ordering.

Definition 5. *Given a relation $\preceq$ on S,*

- *an element $s \in S$ is* well *if for every infinite sequence $\{s_i \in S\}_{i \in \mathbb{N}}$ issuing from $s = s_0$, there exist an ordered pair $j < k$ of natural numbers such that $s_j \preceq s_k$;*
- *the relation $\preceq$ is* well *if every element of S is well.*

Example 1. The relation $\geq$ defined on $T(\mathcal{F})$, the set of trees over the signature $\mathcal{F}$ by $\quad t \geq u \;$ iff $\; |t| \geq_{\mathbb{N}} |u| \quad$ is a quasi-ordering, but is not an ordering since different terms may have equal size. Indeed, $\geq$ is well-founded and $\leq$ is a well quasi-ordering. $\Box$

Remark that relations change to their inverse when talking about being a well-relation, or a well-founded relation. This is not purely formal :

Property 3. Let $\preceq$ be a well quasi-ordering on S. Then $\succeq$ is a well-founded quasi-ordering.

Example 2 of the well-founded subtem relationship shows that the converse is not true in general, although it becomes true for total relations: every well-founded total quasi-ordering is a well quasi-ordering.

Example 2. Let $\mathcal{F} = \{0, f(,)\}$. We consider the sequences of terms $\{t_n\}_n$ and $\{u_n\}_n$ defined as: $t_0 = 0$, $t_{n+1} = f(0, t_n)$ and $u_n = f(f(0, 0), t_n)$. The sequence $\{t_n\}_n$ is increasing for the subterm ordering, while the sequence u_n does not contain a pair of comparable terms.

Property 4. Let $\geq$ and $\succeq$ be two quasi-orderings on S such that

1. $\succ \subseteq >$: then, $\succeq$ is a well-founded quasi-ordering if $\geq$ is well-founded quasi-ordering.
2. $\preceq \subseteq \leq$: then $\leq$ is a well quasi-ordering if $\preceq$ is a well quasi-ordering.

Well-founded quasi-orderings are therefore closed under containment while well quasi-orderings are closed under inclusion. This is one reason for their fame. The second is that they can help building well-founded orderings by using property 3.

3 Extension orderings on tuples

We extend a relation from a set to tuples of elements of that set, while preserving the properties of transitivity, wellness and well-foundedness.

3.1 Product extension

Let $(D_1, \succeq_1), \ldots, (D_n, \succeq_n)$ be pairs of a non-empty set together with a relation on that set. The product set $D_1 \times \ldots \times D_n$ of n-tuples of elements of $D_1, \ldots, D_n$ is endowed with the relation $(\succeq_1, \ldots, \succeq_n)_\times$, written for short as $\succeq_\times$ and called *product*, defined as:

$$(a_1, \ldots, a_n) \succeq_\times (b_1, \ldots, b_n) \text{ iff } \forall i \in [1..n]\ a_i \succeq_i b_i$$

$$(a_1, \ldots, a_n) \simeq_\times (b_1, \ldots, b_n) \text{ iff } \forall i \in [1..n]\ a_i \simeq_i b_i$$

Remark that the strict part of the product $\succeq_\times$ implies the strict part of at least one component relation.

Proposition 1. *The relation $\succeq_\times$ is*

- *a quasi-ordering on $D_1 \times \ldots \times D_n$ whose equivalence is the relation $\simeq_\times$ if the relations $\succeq_1, \ldots, \succeq_n$ are quasi-orderings whose equivalences are the relations $\simeq_1, \ldots, \simeq_n$;*
- *well-founded if so are all relations $\succeq_1, \ldots, \succeq_n$;*
- *a well-founded quasi-ordering on $D_1 \times \ldots \times D_n$ iff so is $\succeq_i$ on D_i for all i.*
- *a well quasi-ordering on $D_1 \times \ldots \times D_n$ iff so is $\succeq_i$ on D_i for all i.*

3.2 Lexicographic extension

Let $(D_1, \succeq_1), \ldots, (D_n, \succeq_n)$ be pairs of a non-empty set with a relation on that set. The product set $D_1 \times \ldots \times D_n$ of n-tuples of elements of $D_1, \ldots, D_n$ is endowed with the relation $(\geq_1, \ldots, \geq_n)_{lex}$, written for short as $\geq_{lex}$ (or $\geq^{lex}$ when necessary) and called lexicographic extension of the starting relations, defined as

$$(a_1, \ldots, a_n) >_{lex} (b_1, \ldots, b_n) \text{ iff } \exists i \geq 1, \forall j < i, a_j \simeq_j b_j \text{ and } a_i >_i b_i$$
$$(a_1, \ldots, a_n) \simeq_{lex} (b_1, \ldots, b_n) \text{ iff } \forall i \geq 1, a_i \simeq_i b_i.$$

Proposition 2. *The relation $\geq_{lex}$ is*

- *a quasi-ordering on $D_1 \times \ldots \times D_n$ whose equivalence is the relation $\simeq_{lex}$ if the relations $\succeq_1, \ldots, \succeq_n$ are quasi-orderings whose equivalences are the relations $\simeq_1, \ldots, \simeq_n$;*
- *total iff each component relation is total ;*
- *well-founded iff so are all relations $\succeq_1, \ldots, \succeq_n$;*
- *a well-founded quasi-ordering on $D_1 \times \ldots \times D_n$ iff this is true of every relation $\succeq_i$ on D_i ;*
- *a well quasi-ordering on $D_1 \times \ldots \times D_n$ iff son is $\succeq_i$ on D_i for all i.*

Example 3. Given two constants a and b, let $\mathcal{R}$ be the rewrite system

$$\begin{cases} (R1)\ x \cdot (y \cdot z) \rightarrow y \cdot a \\ (R2)\ (x \cdot y) \cdot a \rightarrow x \cdot (y \cdot b) \end{cases}$$

We construct two interpretation functions: $\phi_1(u)$, the size of u, and $\phi_2(u)$, the number of ocurrences of a in u. Let $\phi(u) = (\phi_1(u), \phi_2(u)) \in \mathbb{N}^2$. We use $>$ for the ordering on natural numbers. It is easy to verify that

(i) if u rewrites to v with $R1$, then $\phi_1(u) > \phi_1(v)$.

(ii) if u rewrites to v with $R2$, then $\phi_1(u) = \phi_1(v)$ and $\phi_2(u) > \phi_2(v)$.

Therefore, $\phi(u) >_{lex} \phi(v)$. Since $\geq_{lex}$ is well-founded, $\mathcal{R}$ is terminating.

The product extension is of course (strictly) included into the lexicographic extension. It is nevertheless very useful by ensuring that all components of a tuple decrease in their respective ordering.

4 Extension orderings on multisets

If S is a set, the set $\mathcal{M}(\mathcal{S})$ of finite *multisets* over S is the set of applications from S to $\mathbb{N}$ which take value 0 for all elements but a finite number of them. Usually, we denote a multiset $\mathcal{M}$ by repeating the element d a number of times equal to $\mathcal{M}(d)$.

Example 4. Let $S = \mathbb{N}$ and $\mathcal{M}$ be the multiset $\mathcal{M}(0) = 2$, $\mathcal{M}(1) = 0$, $\mathcal{M}(2) = 3$ et $\mathcal{M}(n) = 0$ for $n > 2$, that is $\mathcal{M} = \{0, 0, 2, 2, 2\}$.

We define on multisets the operations $+$, $|\_|$ and the relation $\in$ such that:

$$\begin{aligned} \forall x, y \in \mathcal{M}(S), \forall d \in S\ (x + y)(d) &= x(d) + y(d) \\ |x| &= \textstyle\sum_{d \in S} x(d) \\ d \in x \text{ iff } x(d) &> 0 \end{aligned}$$

We aim here at building relations on the set of multisets constructed over a set S endowed with some relation, an idea originating from [3]. In the literature, there exist numerous definitions of the multiset extension of a relation $\geq_S$ from the set S to the sets of multisets over S. All coincide for the case in which we are interested, when $\geq_S$ is a quasi-ordering. We give one which is suitable for reasonning. Hackers may favour others.

Definition 6. *The multiset relation on $\mathcal{M}(\mathcal{S})$ generated by a relation $\geq$ defined on S, written as $\geq_{mul}$ (or $\geq^{mul}$ when necessary), is the reflexive, transitive closure of the relation*

$$\geq_{\mathcal{M}(\mathcal{S})} = (>_{\mathcal{M}(\mathcal{S})} \cup \simeq_{\mathcal{M}(\mathcal{S})})$$

where $\begin{cases} M + \{x\} >_{\mathcal{M}(\mathcal{S})} M + \{y_1, \ldots, y_n\} & \text{if } \forall i \in [1..n]\ x >_S y_i \\ M + \{x\} \simeq_{\mathcal{M}(\mathcal{S})} M + \{y\} & \text{if } x \simeq y \end{cases}$

Note that the multiset extension of a relation is monotonic with respect to the operations of multisets:

$$\forall x, y, z \in \mathcal{M}(S),\ x >_{mul} y \Rightarrow x + z >_{mul} y + z$$

Here is the main property of the multiset extension of a relation on a set:

Proposition 3. *The relation $\geq_{mul}$ is*

- *a quasi-ordering over $\mathcal{M}(\mathcal{S})$ whose associated equivalence and strict ordering are respectively $(\simeq_{\mathcal{M}(\mathcal{S})})^*$ and $(>_{\mathcal{M}(\mathcal{S})})^+$ in case $\geq$ is a quasi-ordering over S ;*
- *a total quasi-ordering iff $\geq$ is total ;*
- *an ordering iff $\geq$ has no loop ;*
- *well-founded on $\mathcal{M}(\mathcal{S})$ iff $\geq$ is well-founded on S ;*
- *a well-founded quasi-ordering iff $\geq$ is well-founded ;*
- *a well quasi-ordering of $\mathcal{M}(\mathcal{S})$ iff $\geq$ is a well quasi-ordering of S.*

This last result is old (almost a century) in the particular case where the starting quasi-ordering is the equality over a finite alphabet. The traditional wording of this result known as *Dickson's lemma* is that in any infinite sequence of monomials over a finite number of indeterminates, there exists a pair of monomials of which one divides the other. This result is used for proving the termination of Buchberger's algorithm which computes a Gröbner basis from a finite set of polynomials.

Proofs of well-foundedness and well-orderedness are well known. The more complex second proof, based on Nash-Williams's minimal counter-example sequence argument, uses transitivity in an essential way.

Example 5. Let us consider the single rule $(x * y) * z \to x * (y * z)$. We interpret a term t by the multiset $\phi(t)$ of numbers $|u|$ for every subterm of t of the form $u * v$. If $t[u]_p$ rewrites to $t[v]_p$ at position p, then $\phi(t[u]_p) = \phi(u) + K$ and $\phi(t[v]_p) = \phi(v) + K$. It therefore suffices to prove that $\phi(u) >_m \phi(v)$, which follows from a simple calculation.

5 Extension orderings on words

Given an alphabet A, we denote by A^* the set of finite words on A, by $u = u_1 \dots u_n$ a word of length $|u| = n$, and by ϵ the empty word.

Definition 7. *Given a reflexive relation $\leq$ on A, we define the* subword *relationship generated by $\leq$ on A^* as the smallest relation $\trianglelefteq$ containing the pair (ϵ, ϵ), and satisfying the following two properties:*
(i) $u \trianglelefteq a \cdot v$ if $u \trianglelefteq v$,
(ii) $b \cdot u \trianglelefteq a \cdot v$ if $b \leq a$ and $u \trianglelefteq v$.

The subword relationship can be obtained as the special case of the above definition of $\trianglelefteq$ in which the reflexive relation on A is the identity.

Lemma 1. *The relation $\trianglelefteq$ is*

- *a quasi-ordering if so is $\leq$;*
- *well-founded if so is $\leq$;*
- *a well-founded quasi-ordering on A^* if so is $\leq$ on A ;*
- *a well quasi-ordering on A^* if so is $\leq$ on A (Higman's lemma).*

6 Extension orderings on trees

Trees generalize words. Given an alphabet $\mathcal{F}$ of function symbols with (possibly variable) arities, and a set of variables $\mathcal{X}$, we consider the set

$\mathcal{T}(\mathcal{F}, \mathcal{X})$ of labelled trees (or terms) over $\mathcal{F} \cup \mathcal{X}$. We denote by $\mathcal{T}(\mathcal{F})$ the set of *ground* trees. A substitution is a mapping from $\mathcal{X}$ to $\mathcal{T}(\mathcal{F}, \mathcal{X})$ extended as an endomorphism from $\mathcal{T}(\mathcal{F}, \mathcal{X})$ to $\mathcal{T}(\mathcal{F}, \mathcal{X})$. We use postfix notation for the application of a substitution. A substitution is ground if $\forall x \in \mathcal{X}\ x\sigma \in \mathcal{T}(\mathcal{F})$.

We start generalizing the subword relationship, called *embedding* in the context of trees. We will therefore be given a reflexive relation $\leq_{\mathcal{F}}$ on $\mathcal{F}$, called *precedence*. In contrast with the previous section, we will not assume that the precedence relation is a quasi-ordering of $\mathcal{F}$.

6.1 Embedding

Definition 8. *Given a reflexive relation $\leq_{\mathcal{F}}$ on $\mathcal{F}$, we define the* embedding *relationship generated by $\leq_{\mathcal{F}}$ on $\mathcal{T}(\mathcal{F}, \mathcal{X})$ as the smallest relation $\trianglelefteq$ such that, given two terms $u = f(u_1, \ldots, u_m)$ and $v = g(v_1, \ldots, v_n)$, then $u \trianglelefteq v$, iff one of the following two cases holds :*

- *$f \leq_{\mathcal{F}} g$ and there exist an increasing subsequence $j_1 < \ldots < j_m$ of $[1..n]$ such that $\forall i \in [1..m]\ u_i \trianglelefteq v_{j_i}$*
- *$\exists i \in \{1, \ldots, n\}$ such that $u \trianglelefteq v_i$.*

For a term headed by f to be embedded in a term headed by g such that $f \leq_{\mathcal{F}} g$ requires the arity of g to be at least the arity of f. In particular, if $a, b \in \mathcal{F}$ are two constants such that $a \simeq_{\mathcal{F}} b$, then $a \trianglelefteq b$ and $b \trianglelefteq a$.

Example 6. Figure 1 shows two terms $t = f(g(f(a, b)), g(f(b, g(a))))$, $u = f(g(f(a, a)), f(a, f(g(h(h(f(a, b)))), g(f(h(f(b, g(a))), h(g(h(a))))))))$ such that $t \triangleleft u$ for the identity relation on $\mathcal{F}$. The corresponding computation can be done in several ways, one being suggested on the figure.

Lemma 2. *Embedding contains the subterm relationship.*

The embedding relation generated by a precedence relation satisfies:

Theorem 1. *The relation $\trianglelefteq$ is*

- *a quasi-ordering if so is the precedence ;*
- *well-founded if so is the precedence ;*
- *a well-founded quasi-ordering on $\mathcal{T}(\mathcal{F}, \mathcal{X})$ if so is $\leq$;*
- *a well quasi-ordering on $\mathcal{T}(\mathcal{F}, \mathcal{X})$ if so is the precedence, a result known as Kruskal's theorem.*

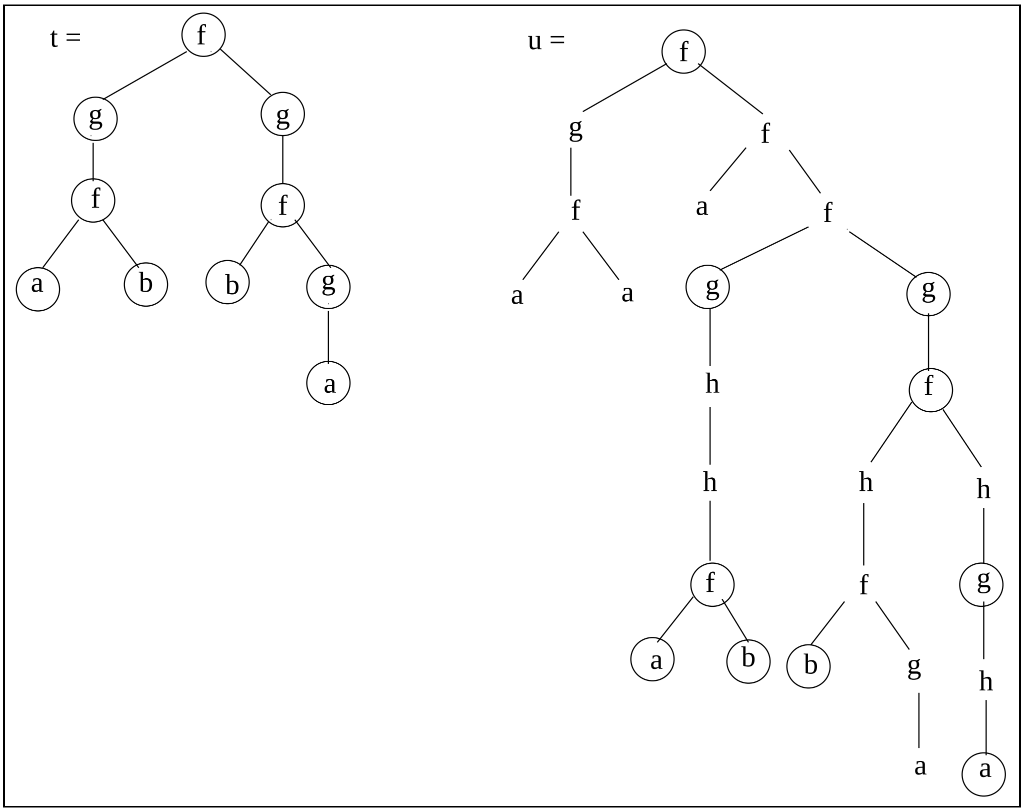

Fig. 1. Example of embedding

The last property is proved by the minimal counter-example sequence argument, invented by Nash-Williams in order to give a (relatively simple) non-combinatorial proof of the special case of this result obtained when the precedence relation is the identity, in which case the generated relation is simply called *embedding*:

Corollary 1. *Any infinite subset of $T(\mathcal{F})$ contains an infinite increasing sequence of trees for the embedding.*

Kruskal's theorem allows a systematic construction of well-founded orderings on terms, called *simplification orderings*:

Definition 9. *A* simplification *quasi-ordering is a quasi-ordering $\geq$ of $T(\mathcal{F}, \mathcal{X})$ whose strict part contains the (strict) subterm relationship and*

(i) is monotonic : $\forall u[]$, $s \geq t \Rightarrow u[s] \geq u[t]$ and $s > t \Rightarrow u[s] > u[t]$,

(ii) is stable : $\forall \gamma$, $s \geq t \Rightarrow s\gamma \geq t\gamma$ and $s > t \Rightarrow s\gamma > t\gamma$.

Proposition 4. *Simplification quasi-orderings are well-founded.*

Proof. It follows from Property 3 using Property 4, since they contain the embedding relationship which is a well-ordering under the finiteness assumption of $\mathcal{F}$. □

Example 7. Many well-founded orderings are simplification orderings: the subterm relationship; the quasi ordering of $T(\mathcal{F}, \mathcal{X})$ defined by comparing the size of terms; the quasi ordering of $T(\mathcal{F}, \mathcal{X})$ defined by terms depth; the lexicographic composition of two simplification orderings.

6.2 Recursive extension on trees

We now come to the recursive path ordering, a simplification ordering containing the embedding. It will therefore be a quasi well-ordering in case of a finite signature, hence well-founded in this case. Moreover, it is known to be well-founded for an infinite signature in case the precedence relation is a well-ordering. We show here that it is enough to assume that the precedence relation is well-founded. The proof follows from [7], in which the more complex case of higher-order terms is considered.

The recursive path ordering builds upon two ingredients:

- a quasi-ordering $\geq_{\mathcal{F}}$ on the signature $\mathcal{F}$;
- a partition of $\mathcal{F}$ into two subsets: one of lexicographic symbols and another of multiset symbols. We say that $f \in Lex$ has the lexicographic *status*, while $f \in Mul$ has the multiset status. The partition $\mathcal{F} = Lex \uplus Mul$ is assumed to satisfy the following properties:
  (i) if $f \simeq_{\mathcal{F}} g$, then f and g have the same status;
  (ii) if $f \simeq_{\mathcal{F}} g \in Lex$, then f and g have the same arity.

Condition (ii) forces varyadic symbols to have a multiset status.

Ground terms We first define the ordering on ground terms:

Definition 10 (Recursive path ordering). *The* recursive path quasi-ordering $\geq_{rpo}$ *extending* $\geq_{\mathcal{F}}$ *is defined as follows:*

$$s = f(\overline{s}) \geq_{rpo} g(\overline{t}) = t \quad \textit{iff}$$

subterm case: $\exists u \in \overline{s},\ u \geq_{rpo} t$
precedence case: $f >_{\mathcal{F}} g$ *and* $\forall v \in \overline{t},\ s >_{rpo} v$
multiset case: $f \simeq_{\mathcal{F}} g$ *and* $f \in Mul$ *and* $\overline{s}\ (\geq_{rpo})_{mul}\ \overline{t}$
lexicographic case: $f \simeq_{\mathcal{F}} g$ *and* $f \in Lex$ *and* $\overline{s}(\geq_{rpo})_{lex}\overline{t}$ *and* $\forall v \in \overline{t}\ s >_{rpo} v$

1. The definition is non-deterministic, in case the subterm case applies;
2. The definition is effective, since recursive calls operate on smaller arguments. More precisely, if $\geq_{rpo}$ is defined for pairs (u, v) whose sizes are strictly smaller than (n, m) for the lexicographic extension of $>_{\mathbb{N}}$, then $\geq_{rpo}^{mul}$ is defined on pairs of multisets of terms $(\{u_1, \ldots, u_p\}, \{v_1, \ldots, v_k\})$ such that, $\forall i \in [1..p]\ \forall j \in [1..k]$, $(|u_i|, |v_j|) <_{\mathbb{N}}^{lex} (n, m)$, and $\geq_{rpo}^{lex}$ is defined on pairs of tuples of terms $(\{u_1, \ldots, u_p\}, \{v_1, \ldots, v_p\})$ such that, $\forall i \in [1..p]$, $(|u_i|, |v_i|) <_{\mathbb{N}}^{lex} (n, m)$.
3. Note the need for lexicographic and multiset extensions of arbitrary relations, since we need to make sure that the definition is by induction before proving that the defined relation is transitive. This important remark is due to Fereira and Zantema [9]. However, there is no need to go through the complications of fixpoint theory as they do.
4. In practice, it is often useful to use a right-to-left lexicographic status. More complex statuses are introduced in [6].

In his original paper introducing the recursive path ordering, Dershowitz considered only the multiset status [4]. The lexicographic status was introduced later by Kamin and Levy in an unpublished note.

Open terms There are two ways to extend the recursive path ordering to open terms. For the first, stability by substitution is the definition:

$$s \succeq_{rpo} t \text{ iff } s\gamma \geq_{rpo} t\gamma \quad \forall\gamma \text{ ground substitution whose domain contains } \mathcal{V}ar(s,t).$$

It was not clear that this definition could be decidable until this was proven by Comon as a consequence of a more general result [2]. It is even NP-complete for simple conditions on the precedence relation. This complex definition can be approximated by one of a quadratic complexity. This second definition is indeed the same as before except that variables are considered as new constants incomparable between themselves (except by reflexivity) and with all function symbols in $\mathcal{F}$ [4]:

$$s \succeq_{rpo} t \text{ iff } s \geq_{rpo} t \text{ with } \geq_{\mathcal{F}\cup\mathcal{X}} = \geq_{\mathcal{F}} \cup \{x \geq x \mid x \in \mathcal{X}\}.$$

From now on, we adopt this definition for the relation $\succeq_{rpo}$.

Example 8. Assume that $\mathcal{F}$ contains two binary symbols $* >_{\mathcal{F}} +$ with multiset status. Then, $(x + y) * z \succeq_{rpo} (x * y) + (x * z)$ since

– $* >_{\mathcal{F}} +$ and $(x + y) * z \succ_{rpo} x * y$ and $(x + y) * z \succ_{rpo} x * z$.
– $x + y \succ_{rpo} x$ and $x + y \succ_{rpo} y$, hence $\{x + y, z\}(\succ_{rpo})_{mul}\{x, y\}$. And therefore $(x + y) * z \succ_{rpo} x * y$. Similarly, $(x + y) * z \succ_{rpo} x * z$.

Properties of the recursive path ordering We prove a succession of properties, whose order aims at making the proofs as simple, as elegant and as independent of each other as possible. They differ from [4] substantially.

Lemma 3. *Assume that* $s \succeq_{rpo} t$ *and* $t = g(\overline{t})$. *Then,* $\forall v \in \overline{t}$, $s \succ_{rpo} v$.

Proof. Let $s \succeq_{rpo} t$. We prove that $s \succ_{rpo} v$ for every $v \in \overline{t}$ by induction on $|s| + |t|$. We then distinguish several cases, whether the proof of $s \succeq_{rpo} t$ is made first by one of the subterm, precedence, lexicographic or multiset cases. We do the subterm case, the others being similar:

Assume that $s \succeq_{rpo} t$ by the subterm case. Then $s = f(\overline{s})$ and $u \succeq_{rpo} t$ for some $u \in \overline{s}$. By induction hypothesis, $u \succ_{rpo} v$, and therefore $s \succeq_{rpo} v$ by the subterm case. Assume that $v \succeq_{rpo} s$. By induction hypothesis again, $v \succ_{rpo} u$, a contradiction. Therefore, $s \succ_{rpo} v$, we are done. □

Corollary 2. *Let* $s \succeq_{rpo} t$ *by subterm or precedence case. Then,* $s \succ_{rpo} t$.

Proof. By contradiction and lemma 3. □

Lemma 4. $s \succeq_{rpo} t$ *and* $t \succeq_{rpo} s$ *iff* $s =_{Mul} t$, *where*

$$s =_{Mul} t \textit{ iff } s = f(\overline{s}), t = f(\overline{t}), f \simeq_{\mathcal{F}} g \in Mul \textit{ and } \overline{s}(=_{mul})_{mul}\overline{t}$$

Proof. The proof of the only if case is by induction on $|s|$. □

Corollary 3 (reflexivity). $\succeq_{rpo}$ *is reflexive.*

Corollary 4. $\succ_{rpo}$ *has the (strict) subterm property.*

Proof. By using Corollary 3 and Lemma 3. □

Lemma 5 (transitivity). *Assume that* $s \succeq_{rpo} t \succeq_{rpo} u$. *Then* $s \succeq_{rpo} u$, *with* $s \succ_{rpo} u$ *iff* $s \succ_{rpo} t$ *or* $t \succ_{rpo} u$.

Proof. Assume that $s = f(\overline{s}) \succeq_{rpo} t = g(\overline{t}) \succeq_{rpo} u = h(\overline{u})$. The proof is by induction on $|s| + |t| + |u|$ and by cases. We write (a, b) to indicate that $s \succeq_{rpo} t$ is by case a and $t \succeq_{rpo} u$ by case b.

1. case (subterm, any). Then $s_i \succeq_{rpo} t \succeq_{rpo} u$ for some s_i. By the induction hypothesis, $s_i \succeq_{rpo} u$, hence $s \succ_{rpo} u$ by subterm.
2. case ($\neq$ subterm, subterm). Then $t_i \succeq_{rpo} u$. By lemma 3, $s \succ_{rpo} t_i$, hence $s \succ_{rpo} u$ by the induction hypothesis.
3. ($\neq$ subterm, precedence). Then $f \geq_{\mathcal{F}} g >_{\mathcal{F}} h$ and $t \succ_{rpo} u_i$ for each u_i. By the induction hypothesis, $s \succ_{rpo} u_i$ for each u_i and $s \succ_{rpo} t$ by the precedence case (using Corollary 4).

4. case (precedence, multiset or lexicographic). Then $t \succ_{rpo} u_i$ for every u_i by lemma 3. By the induction hypothesis, $s \succ_{rpo} u_i$ for every u_i and $s \succeq_{rpo} u$ by precedence case (using Corollary 4).
5. case (multiset or lexicographic, multiset or lexicographic). By the induction hypothesis. $\square$

Lemma 6. *Assume that $\geq_{\mathcal{F}}$ is total on $\mathcal{F}$. Then $\succeq_{rpo}$ is a total quasi-ordering on ground terms.*

Proof. By corollary 3 and lemma 5, we know that $\succeq_{rpo}$ is a quasi-ordering. We show totality by contradiction. Assume that $s = f(\overline{s}), t = g(\overline{t})$ is the smallest pair of ground terms such that s and t are incomparable. By the minimality assumption, all pairs $(s, t_j), (s_i, t), (s_i, t_j)$ are comparable. If $t_j \succeq_{rpo} s$ for some $t_j \in \overline{t}$, then $t \succ_{rpo} s$, a contradiction. Hence $s \succ t_j$ for all $t_j \in \overline{t}$. Similarly, $t \succ s_i$ for all $s_i \in \overline{s}$. We distinguish:

1. $f >_{\mathcal{F}} g$. Then $t \succ_{rpo} s$ by the precedence case, a contradiction.
2. $g >_{\mathcal{F}} f$ is similar.
3. $f = g \in Mul$. Since all pairs (s_i, t_j) are comparable, either $\overline{s}(\succeq_{rpo})_{mul}\overline{t}$, in which case $s \succeq_{rpo} t$; or $\overline{t}(\succeq_{rpo})_{mul}\overline{s}$, in which case $t \succeq_{rpo} s$. Both cases yield a contradiction.
4. $f = g \in Lex$. Again, $\overline{s} \succ_{mul} \overline{t}$, with $s \succeq_{rpo} t$ since $s \succ_{rpo} t_j$ for all $t_j \in \overline{t}$, or else $\overline{t} \succ_{mul} \overline{s}$, with $t \succeq_{rpo} s$ since $t \succ_{rpo} s_i$ for all $s_i \in \overline{s}$, yielding again a contradiction in both cases. $\square$

Lemma 7. *$\succ$ is monotonic.*

Proof. Follows easily from the multiset or lexicographic cases and lemma 3. $\square$

Lemma 8. *$\succ$ and $\succeq$ are stable relations.*

Proof. We only show the non-obvious case of the strict relation.

Let $s = f(s_1, \ldots, s_n) \succ_{rpo} t = g(s_1, \ldots, s_n)$, with $f, g \in (\mathcal{F} \cup \mathcal{X})$, and σ be a substitution. We show that $s\gamma \succ_{rpo} t\gamma$ by induction on $|s|+|t|$. All four cases (depending upon the proof of $s \succ_{rpo} t$) are easy. $\square$

Well-orderedness and well-foundedness It follows from the previous results:

Property 5. $\geq_{rpo}$ is a simplification quasi-ordering.

Theorem 2. *$\leq_{rpo}$ is a well-ordering on the set of terms when the precedence relation is a well-ordering on $\mathcal{F}$.*

As a corollary, $\geq_{rpo}$ is a well-founded quasi-ordering when the precedence relation is a well-ordering of $\mathcal{F}$. We will now show that $\geq_{rpo}$ is a well-founded quasi-ordering when the precedence relation is a well-founded quasi-ordering, a result stronger than the previous corollary. The proof does not use Kruskal's theorem, and is indeed elementary.

Lemma 9. *Let $f \in \mathcal{F}$ have arity n, and let $\overline{s}$ be a vector of n terms strongly normalizing for $\succ_{rpo}$. Then $f(\overline{s})$ is strongly normalising.*

Proof. The idea is to consider the set T of terms which are smaller than the terms in $\overline{s}$ for $\succ_{rpo}$. The restriction of $\succ_{rpo}$ to terms in T is a well-founded ordering which will now be used for building an outer induction on the pairs $(f, \overline{s})$ ordered by $(>_{\mathcal{F}}, (\succ_{rpo})_{stat})_{lex}$, where $stat$ indicates the status (multiset or lexicographic) of f. This ordering is well-founded, since it is built from well-founded orderings by using mappings that preserve well-founded orderings.

We now prove that $f(\overline{s})$ is strongly normalizing by proving that for all t such that $f(\overline{s}) \succ_{rpo} t$, then t is strongly normalizing. This property is itself proved by an (inner) induction on $|t|$, and by case analysis upon the proof that $f(\overline{s}) \succ_{rpo} t$.

1. subterm case: $\exists u \in \overline{s}$ such that $u \succ_{rpo} t$. By assumption, u is strongly normalizing, hence so is its reduct t.
2. precedence case: $t = g(\overline{t})$, $f >_{\mathcal{F}} g$, and $\forall v \in \overline{t}$, $s \succ_{rpo} v$. By inner induction, v is strongly normalizing, hence so is $\overline{t}$. By outer induction, $g(\overline{t}) = t$ is strongly normalizing.
3. multiset case: $t = f(\overline{t})$ with $f \in Mul$, and $\overline{s}(\succ_{rpo})_{mul}\overline{t}$. By definition of the multiset extension, $\forall v \in \overline{t},\ \exists u \in \overline{s}$ such that $u \succeq_{rpo} v$. Since $\overline{s}$ is a vector of strongly normalizing terms by assumption, so is $\overline{t}$. We conclude by outer induction that $f(\overline{t}) = t$ is strongly normalizing.
4. lexicographic case: $t = f(\overline{t})$ with $f \in Lex$, $\overline{s}(\succ_{rpo})_{lex}\overline{t}$, and $\forall v \in \overline{t},\ s \succ_{rpo} v$. By inner induction, $\overline{t}$ is strongly normalizing, and by outer induction, so is $f(\overline{t}) = t$. □

Theorem 3. *$\succ_{rpo}$ is well-founded.*

Proof. We show that every term t is strongly normalizing by induction on $\mid t \mid$. Let $t = f(\overline{t})$. By induction hypothesis, $\overline{t}$ is strongly normalizing, and by lemma 9, so is t. □

Besides its conceptual beauty, Dershowitz's recursive path ordering provides with a very concise ordinal notation. For example, assuming that $+$ is binary and that $+ >_{\mathcal{F}} 0$, where 0 is a constant, the term $(((0 + 0) + 0) + 0) + 0$ corresponds to ϵ_0.

7 Conclusion

We believe that our proof method has an important potential for improving or proving well-foundedness properties of extension orderings.

The recursive path ordering generalizes to flattened trees, by using variable arity symbols when they are associative and commutative. This generalization required a series of papers culminating with [8]. It would be interesting to apply our proof method to this case. The recursive path ordering also generalizes to the case of simply typed higher-order languages: the *higher-order recursive path ordering* of Jouannaud and Rubio is is again well-founded, a property whose proof generalizes the one given here for the recursive path ordering by using *reducibility interpretations* of types *à la Tait and Girard*. This result was in turn generalized to the whole calculus of constructions by Walukiewicz [10].

Acknowledgments: we are grateful to H. Comon and A. Rubio for sharing ideas, and to J. Meseguer for carefully reading the manuscript.

References

1. Hubert Comon. Solving inequations in term algebras. In *Proc. 5th IEEE Symp. Logic in Computer Science, Philadelphia*, June 1990.
2. Hubert Comon. Solving symbolic ordering constraints. *International Journal of Foundations of Computer Science*, 1(4):387–411, 1990.
3. Nachum Dershowitz and Zohar Manna. Proving termination with multiset orderings. *Communications of the ACM*, 22(8):465–476, August 1979.
4. Nachum Dershowitz. Orderings for term rewriting systems. *Theoretical Computer Science*, 17(3):279–301, March 1982.
5. Nachum Dershowitz and Jean-Pierre Jouannaud. Rewrite systems. In J. van Leeuwen, editor, *Handbook of Theoretical Computer Science*, volume B, pages 243–309. North-Holland, 1990.
6. Maribel Fernández and Jean-Pierre Jouannaud. Modular termination of term rewriting systems revisited. In Egidio Astesiano, Gianni Reggio, and Andrzej Tarlecki, editors, *Recent Trends in Data Type Specification*, LNCS 906:255–272. Springer-Verlag, 1995.
7. Jean-Pierre Jouannaud and Albert Rubio. The higher-order recursive path ordering. In *Proc. 14th Annual IEEE Symposium on Logic in Computer Science*, Trento, Italy. IEEE Comp. Soc. Press.
8. Albert Rubio. A fully syntactic AC-RPO. In Narendran and Rusinowitch eds, Proceedings of 10th RTA, Lecture Notes in Computer Science 1631:133–147, 1999.
9. Hans Zantema. Termination. In *Term Rewriting Systems*, Marc Bezem, Jan Wilhelm Klop and Roel de Vrijer eds, Cambridge Tracts in Theoretical Computer Science 55, Cambridge U. Press, 2003.
10. Daria Walukiewicz-Chrzaszcz. Termination of rewriting in the Calculus of Constructions. In *Proc. of the Workshop on Logical Frameworks and Meta-languages, Santa Barbara, California*, 2000.