J. Symbolic Computation (2000) **29**, 921–957
doi:10.1006/jsco.1999.0333
Available online at http://www.idealibrary.com on IDEAL

Invariants, Patterns and Weights for Ordering Terms

URSULA MARTIN AND DUNCAN SHAND[†]

School of Computer Science, University of St Andrews, St Andrews, Fife KY16 9SS, U.K.

We prove that any simplification order over arbitrary terms is an extension of an order by weight, by considering a related monadic term algebra called the spine. We show that any total ground-stable simplification order on the spine lifts to an order on the full term algebra. Conversely, under certain restrictions, a simplification ordering on the term algebra defines a weight function on the spine, which in turn can be lifted to a weight order on the original ground terms which contains the original order. We investigate the Knuth–Bendix and polynomial orders in this light. We provide a general framework for ordering terms by counting embedded patterns, which gives rise to many new orderings. We examine the recursive path order in this context.

1. Introduction

Orders on terms and other structures have been much studied in the context of termination proofs of rewriting systems, algorithms, logic programs and the like, and since the pioneering work of, for example, Dershowitz (1982), Lescanne (1981) and others an extraordinary variety of orders has been found: see Steinbach (1995) for a recent survey. More recently attention has focused on how we might classify the orders of interest, and as a first step compute numerical or logical invariants for them. This is valuable not only for understanding more clearly the apparent diversity of these complicated structures, but also because it may help us see more readily how to try to prove termination. Furthermore, in some cases these invariants give us more information about the termination problems we are trying to solve.

We call a topological space C a classifying space for a set D if there is a mapping from D to C: thus if elements of D have different images under the mapping they must be distinct. An invariant for an element of D is some number or ordinal naturally associated with elements of C, and hence, under the mapping, with D, so that if elements of D have distinct invariants they are distinct. Thus, for example, square matrices in one of the usual normal forms form a classifying space for linear transformations of a vector space, and the rank or the determinant are numerical invariants.

Every well-founded total order is order-isomorphic to an ordinal called its order type, which may be regarded as measuring its proof theoretic strength. If a well-order $\succ$ proves that a rewrite system $\mathcal{R}$ is terminating, then the order-type of the order can be related to the derivation complexity of $\mathcal{R}$ (Hofbauer, 1992; Cichon and Weiermann, 1997), which in turn can be related to the proof theoretic and algorithmic complexities of the relation

[†]E-mail: {um,ddshand}@dcs.st-and.ac.uk

0747–7171/00/060921 + 37 $35.00/0

being computed by $\mathcal{R}$ (Cichon and Weiermann, 1997). For total monotonic orders over terms in unary function symbols the order types have been classified (Hofbauer, 1992; Touzet, 1997): for example, for terms over two unary function symbols, there are three possible order types (Martin and Scott, 1997).

Thus the order type is a very weak invariant. Scott (1994) went much further and showed that any total monotonic order on terms in unary function symbols can be linearized: that is, it is an extension of an order by weight, defined by first weighting each of the function symbols with a non-negative weight, and then weighting terms by adding up the weights of the function symbols they contain. Any pre-order by weight can be extended in continuum many ways to a total order of this kind, and different weightings give different orders provided one is not a scalar multiple of the other: in the case of two function symbols Martin and Scott (1997) and Prohle and Perlo-Freeman (1997) provide a finer classification.

It follows that any such order $\succ$ over Σ, a set of n unary function symbols, is associated with an equivalence class $p(\succ) = [(w_1, \ldots, w_n)] = \{(aw_1, \ldots, aw_n) \mid 0 < a \in \mathbb{R}\}$ of sequences of non-negative real weights. The set of all such $p(\succ)$ bijects with $\mathcal{P}_n$, the non-negative orthant of projective n-space, and this establishes $\mathcal{P}_n$ as a classifying space for these orders. This means that there is a surjective mapping from the set of orders to the classifying space, which is a topological space, and we may use this to regard the set of orders as a topological space also. Further consideration of the weights allows us to construct various invariants.

If $\mathcal{R}$ is a set of rewrite rules over Σ^* we may define a subset $\mathcal{R}_>$ of $\mathcal{P}_n$, namely the subset (possibly empty) of all those p for which any ordering $\succ$ with $p = p(\succ)$ proves $\mathcal{R}$ terminating. Thus, for example for $\mathcal{R} = \{a^2 \longrightarrow b^3\}$ we have $\mathcal{R}_> = \{[(\alpha, \beta)] \mid 2\alpha > 3\beta\}$, which is an open connected subset of $\mathcal{P}_n$.

The classifying space and the associated invariants have been applied by Martin (1996) to classify the rewriting systems they prove terminating: these are group and semigroup presentations. In a similar vein Martin (1989), Mora and Robbiano (1986) and others classified monotonic orders on multisets, and this classification has been used to investigate (Faugere *et al.*, 1993) the so-called Gröbner walk and Gröbner fan of an ideal in a polynomial ring. This enables us to understand all Gröbner bases for an ideal in a uniform framework, and to transform one into another without recomputing from scratch. For both groups and Gröbner bases this has important practical applications, in that a rewrite system which is fast to produce (for example because it has few rules) may not be fast to compute with (for example because it has long derivation sequences) (Linton and Shand, 1996).

The above results concern numerical invariants and classifying spaces for orders on strings and multisets. This paper is concerned with developing similar results for terms. The first main section considers how we may assign numerical invariants to orders on terms, and hence establish $\mathcal{P}_n$ as a classifying space for term orders over n non-constant function symbols. The second concerns a general framework for ordering terms by counting embedded patterns: we construct a large class of new term orders and show how our method subsumes earlier constructions. A final section looks at the recursive path order in the light of our results. We now explain our results in more detail.

1.1. LINEARIZABLE ORDERS

The first section of this paper considers how we may associate numerical invariants to term orders. That is to say, we establish a uniform framework for classifying and comparing orders on terms, with the aim of finding some analogue of Scott's results for strings. At first sight this might seem somewhat implausible: there are an enormous number of monotonic orders on terms, and it might seem unlikely that we could come up with anything as elegant as Scott's result. The key is to identify a distinguished subset of the term algebra on which the orders we are considering behave like orders on strings, then to apply Scott's methods to this subset, and finally to relate the results to the original order.

As a running example throughout this section we consider the polynomial order on $T = T(\{f, g, x, y, d\})$, for f, g binary operators, d a constant and x, y variables. We define $[x] = X, [y] = Y, [d] = D, [f(x, y)] = 2XY^2$ and $[g(x, y)] = X^2 + Y^2$.

Our distinguished subset S, called the spine of T, will be defined as follows in this case, and analogously in general. Let $F(x) = f(x, x)$, $F^{i+1}(x) = f(F^i(x), F^i(x))$ and $G(x) = g(x, x)$, $G^{i+1}(x) = g(G^i(x), G^i(x))$. Then the spine of T is $S = T(\{F, G, x\})$, the set of terms on the unary operators F, G. Thus in our example the spine contains $g(f(x, x), f(x, x)) = G(F(x))$ and $f(g(x, x), g(x, x)) = F(G(x))$.

If $\succ$ is any order on T which is a total stable simplification order on restriction to S, then by Scott (1994) there exists τ with $\infty \geq \tau \geq 0$ such that for all i, j we have

$$F^i(x) \succ G^j(x) \qquad \text{if } i > j\tau \qquad \text{and} \qquad F^i(x) \prec G^j(x) \qquad \text{if } i < j\tau.$$

Further if we define $\mu : S \to \mathbb{R}$ by

$$\mu(s) = \sum_{p=1}^{k} (i_p + \tau\, j_p)$$

for

$$s = F^{i_1} G^{j_1} F^{i_2} G^{j_2} \ldots F^{i_k} G^{j_k}(x),$$

then $s \succ t$ if $\mu(s) > \mu(t)$ and $s \prec t$ if $\mu(s) < \mu(t)$. The sequence $(1, \tau)$ is a numerical invariant of the order $\succ$ on T: in general over n symbols there is a sequence of n values called the weight sequence which behaves similarly. The equivalence class of the weight sequence under scalar multiplication is an element of $\mathcal{P}_n$, establishing this as a classifying space for our class of orders.

For our running example we have

$$\begin{array}{rclcl}
[G(G(x))] & = & [g(g(x,x), g(x,x))] & = & 8X^4 \\
 & & [g(g(x,x), f(x,x))] & = & 4X^6 + 4X^4 \\
[G(F(x))] & = & [g(f(x,x), f(x,x))] & = & 8X^6 \\
 & & [f(g(x, g(x,x)), x)] & = & 8X^6 + 2X^4 \\
[F(G(x))] & = & [f(g(x,x), g(x,x))] & = & 16X^6 \\
[F(F(x))] & = & [f(f(x,x), f(x,x))] & = & 16X^9.
\end{array}$$

Thus, for $i, j \geq 1$, $F^i(x) \succ G^j(x)$ if $[F^i(x)] = aX^{3^i}$ is greater than $bX^{2^j} = [G^j(x)]$ in the polynomial ordering, where a, b are certain powers of 2. Thus $\tau = \log 2/\log 3$. The restriction of this polynomial order to S is total: it is just a certain lexicographic extension of the weight order defined by τ. The order $\succ$ corresponds to $[(1, \log 2/\log 3)]$, the class of $(1, \log 2/\log 3)$ in $\mathcal{P}_2$.

Now consider the rewrite system $\mathcal{R} = \{F(F(x)) \longrightarrow G(G(G(x)))\}$, and the subset $\mathcal{R}_>$ defined above. Our polynomial ordering lies in $\mathcal{R}_>$, and in fact $\mathcal{R}_> = \{[(1,\alpha)] \,|\, \alpha < 2/3\}$ a connected open subset of $\mathcal{P}_2$.

In general the description we have of the restriction of $\succ$ to S also gives us an alternative way of looking at $\succ$ on T. For it turns out that every element u of T has a least upper bound $\bar{u}$ in S, and if $\bar{u} \prec \bar{v}$, then $u \not\succ v$. Thus in our example $u = g(g(x,x), f(x,x))$ has a least upper bound $\bar{u} = G(F(x))$ in S and $v = f(g(x, g(x,x)), x)$ has a least upper bound $\bar{v} = F(G(x))$: the table above confirms that $u \prec \bar{u} \prec v \prec \bar{v}$.

We can go further: if we have a total simplification order on S we can lift it under certain conditions to a simplification pre-order $\trianglerighteq$ on T. If the order on S is the restriction to S of some ordering $\succ$ on T, then $\succ \subseteq \trianglerighteq$. Thus in our running example we can lift the weight-lexicographic order on S induced by the polynomial order to a pre-order on T which contains the polynomial order.

In Section 3 we develop this theory precisely. The subalgebra generated by $\{F, G\}$ is an example of a so-called one-parameter family in T, that is, a subalgebra that is isomorphic to a free algebra on unary function symbols. An order $\succ$ is said to be linearizable over a one-parameter family S if there is a weight sequence (μ_i) satisfying the appropriate generalization of the property of the example. As $\succ$ is total on $S = T(\mathcal{G})$, we may assume without loss of generality that

$$\mathcal{G} = \{{}_1G, \ldots, {}_nG\}$$

where ${}_nG(x) \succ {}_{n-1}G(x) \succ \cdots \succ {}_1G(x)$. Then $\succ|_S$ is linearizable over S if there is a sequence of real numbers, called a weight sequence, satisfying

$$1 = \mu_n \geq \mu_{n-1} \geq \cdots \geq \mu_1 \geq 0$$

such that $\succ|_S \supseteq >_\mu$, where the weight function $\mu : S \to \mathbb{R}$ is defined by $\mu(s) = \sum_{{}_iG \in \mathcal{G}} \#({}_iG, s)\mu_i$ and $s >_\mu t$ if and only if $\mu(s) > \mu(t)$.

In our running example $F = {}_2G, G = {}_1G$ and $1 = \mu_2 > \mu_1 = \log 2/\log 3$ so for example $\mu(F(G(x))) = 1 + \log 2/\log 3$.

Our main theorems are as follows.

THEOREM 3.3. *Let T be a term algebra, $S = T(\mathcal{G})$ a one-parameter family in T and $\succ$ a σ-stable simplification order on T which is total on restriction to S. Let $u(x), v(x)$ be elements of S. Then there exists a τ $(\infty \geq \tau \geq 0)$ such that*

$$u^i(x) \succ v^j(x) \quad \text{if } i > j\tau \text{ and}$$
$$u^i(x) \prec v^j(x) \quad \text{if } i < j\tau.$$

THEOREM 3.4. *Let T be a term algebra, $S = T(\mathcal{G})$ a one-parameter family in T and $\succ$ a σ-stable simplification order on T which is total on S. Then $\succ|_S$ is linearizable over S with classifying space $\mathcal{P}_n$.*

THEOREM 3.5. *Let T be a term algebra, $S = T(\mathcal{G})$ a one-parameter family in T, where $\mathcal{G} = \{{}_1G(x), \ldots, {}_nG(x)\}$, and $\succ$ a σ-stable simplification order on T which is total on S. Then for each s in S there exist non-negative integers $i_1, \ldots, i_n$ and permutations π, ψ of $\{1, 2, \ldots, n\}$ such that*

$${}_{\pi(1)}G^{i_{\pi(1)}}\,{}_{\pi(2)}G^{i_{\pi(2)}} \cdots {}_{\pi(n)}G^{i_{\pi(n)}}(x) \succeq s \succeq {}_{\psi(1)}G^{i_{\psi(1)}}\,{}_{\psi(2)}G^{i_{\psi(2)}} \cdots {}_{\psi(n)}G^{i_{\psi(n)}}(x).$$

The spine, which we described informally above, is a particular example of a one-parameter family, having additional properties which allow us to lift orders on the spine back to T. We have two main results of this kind:

THEOREM 3.6. *Let T be a term algebra containing constants, and let $\succ$ be a total simplification order on the spine S of T, preserved under substitutions $\sigma : \{x\} \to S$. Then there is a simplification pre-order $\trianglerighteq$ on T such that if $s, t \in S$ and $s \succ t$, then $s\sigma \trianglerighteq t\sigma$ for any ground substitution σ.*

THEOREM 3.7. *Let $\succ$ be an order on T which is a total simplification order on restriction to S, preserved under ground substitutions $\sigma : \{x\} \to T$, with weight function μ on S. Suppose that there is a unique maximal constant. Then there is a function $\bar{\mu} : T \to \mathbb{R}$ such that $\succ \subseteq \geq_{\bar{\mu}}$ so that if $u \succ v$, then $\bar{\mu}(u) > \bar{\mu}(v)$.*

In our example we have $\bar{\mu}(g(g(x,x), f(x,x)) = 1 + \log 2/\log 3$.

We give further examples of one-parameter families in Section 4.1, and compute weight sequences for some of the standard orders in Sections 4.2 and 4.3.

The Knuth–Bendix order is a particular order by weight defined on terms: we show that any weight sequence arises as the weight sequence of the restriction of a Knuth–Bendix order to a suitable one-parameter family.

THEOREM 4.7. *Let $0 < \mu_1 \leq \cdots \leq \mu_n = 1$ and $T = T(\mathcal{F} \cup \mathcal{X})$ where $|\mathcal{F} \setminus \mathcal{F}_C| = n$ and $|\mathcal{F}_C| \geq 1$. Then there is a $\mathcal{G} \subset T$ with $|\mathcal{G}| = n$, a one-parameter family $S = T(\mathcal{G} \cup \{x\})$ and a simplification order $>$ on T which is total on restriction to S and has weight sequence $(\mu_1, \ldots, \mu_n)$ over S.*

We then investigate general Knuth–Bendix orders, computing the weight sequences and, in the case where $|\mathcal{G}| = 2$, the further invariants of Martin and Scott (1997). The polynomial orders, first defined in Lankford (1975) and refined by, for example, Ben Cherifa and Lescanne (1987), give us a rich source of examples, and in Section 4.3 we use a technique of Cropper and Martin (2000) to compute the weight sequences for polynomial orders. In particular, if one of the generators of S has a non-linear interpretation, then the weight sequence of S is determined solely by the degrees of the interpretations of the generators of S.

1.2. ORDERS BY COUNTING PATTERNS

The results above concern ordering terms inside one-parameter families by counting occurrences of embedded subterms of the form $F^k(x)$ for suitable patterns F. Orders by counting occurrences of various kinds of patterns, which may be subterms or some other combinatorial device, are a recurrent theme in the theory of orders, beginning with Knuth and Bendix's original paper on rewriting (Knuth and Bendix, 1970) where they described how to order terms by extending a weight function on the function symbols, and continuing with various developments by Steinbach (1995), Martin (1993) and others.

The principle of such constructions over a term algebra T is to count occurrences of some set of patterns P_1, to order the vectors thus obtained using a division order $>_1$ on vectors, to lift this order to obtain a pre-order on T, and then to break ties by considering a new set of patterns P_2. Certain compatibility conditions on the P_i are needed to ensure that a simplification order is obtained. In Section 5 we present a very general framework for such constructions, which allows us to define continuum many new monotonic orders, even over a fixed signature. Essentially we count embedded subterms which are members of a pattern class, a set of variable arity terms P which is closed under taking principal subterms. The compatibility conditions are expressed by requiring that P is the disjoint union of subsets P_i, where each P_i (for $i \geq 2$) does not contain a term and any of its principal subterms, together with certain conditions on the $>_i$.

As an example consider the two terms $s = g(a, g(a, g(a, a)))$ and $t = g(g(a, a), g(a, a))$. Both contain the same number of occurrences of the varyadic subterms g, a and $g(a, a)$. However, s contains more occurrences of the varyadic subterm $g(g)$ than t: our methods show that this allows us to construct a monotonic order $\succ$ with $s \succ t$, using the pattern class $\{a, g, g(g)\}$. Similarly t contains more occurrences of the varyadic subterm $g(g, g)$ than s, and we may construct a monotonic order $\succ'$ with $t \succ' s$, using the pattern class $\{a, g, g(g, g)\}$.

Section 5.1 presents the main result: proving that our construction gives rise to a monotonic order. In Section 5.2 we show the power of the new orders by showing that uncountably many distinct orders may arise in this way, even for two unary function symbols. Specifically we show that for any positive real λ there is an order $\succ_{(\lambda)}$ with $f^N g^{2M} f^N \succ_{(\lambda)} g^M f^{2N} g^M$ for $N/M > \lambda$ and $g^M f^{2N} g^M \succ_{(\lambda)} f^N g^{2M} f^N$ for $N/M < \lambda$. So for example if $\lambda = \pi$, then

$$f^{32} g^{20} f^{32} \succ_{(\lambda)} g^{10} f^{64} g^{10} \succ_{(\lambda)} f^{31} g^{20} f^{31}.$$

A suitable pattern class in this case is $\{f, g, fg, gf, fgf, gfg\}$. The order is induced by assigning the weight $\alpha + \beta\lambda$ to a term containing α occurrences of fgf and β occurrences of gfg and comparing the weights.

We then show that the weight part of the Knuth–Bendix order (Knuth and Bendix, 1970), and Martin's "zig-zag" order (Martin, 1993) are special cases of our new construction.

1.3. PATTERNS FOR SYNTACTIC ORDERS

We continue our investigations of patterns by considering in Section 6 the recursive path order, $>_{\text{rpo}}$ which can be expressed using numeric invariants but does not quite fit the construction of Section 5. We show first that it cannot be regarded as an extension of any order constructed using the technique of Section 5. However, it does correspond to a pre-order on another numeric invariant: the height of a maximal subterm in the largest operator. If $\mathcal{F}$ is a set of function symbols equipped with a precedence $\succ$, and f in $\mathcal{F}$ is not a constant and is larger than any other element of $\mathcal{F}$ under $\succ$, we define skel(s), the f-skeleton of a term s, to be a maximal embedded subterm of s involving only f and constants. We prove that $>_{\text{rpo}}$ extends the ordering induced by the heights of the skeletons, that is to say:

THEOREM 6.7. *Let $f, \mathcal{F}$ and $\succ$ be as above. Then, for $s, t \in T(\mathcal{F})$,*

(1) If $\text{height}(\text{skel}(s)) > \text{height}(\text{skel}(t))$*, then* $s >_{\text{rpo}} t$
(2) If $s >_{\text{rpo}} t$*, then* $\text{height}(\text{skel}(s)) \geq \text{height}(\text{skel}(t))$.

This result is true in particular when $\mathcal{F}$ contains only one non-constant function symbol, bigger than all constants, and reduces to the statement that in this case the recursive path order is just an extension of the order by height. Our next result, Theorem 6.10, allows us to say more, by showing that in this case terms of equal height are ordered by considering certain subterms of maximal depth. This may seem a specialized class, but it includes for example all combinator terms on the binary function symbol. (explicit application), and constants K, I and Y.

It is tempting to conjecture that Theorem 6.7 generalizes, that is that for arbitrary signatures the recursive path order extends the recursive path order on f-skeletons, but we give an example to show that this is false. However, for the case of two unary function symbols Theorem 6.12 shows that a result of this kind is true: it involves counting not arbitrary subterms but subterms which are maximal in a certain class, a technique it may be possible to generalize.

2. Definitions

2.1. Basics

The set of natural numbers $\{0, 1, 2, \ldots\}$ is denoted $\mathbb{N}$ and the set of positive natural numbers $\{1, 2, 3, \ldots\}$ is denoted $\mathbb{N}_+$. The set of real numbers (respectively positive real numbers) is denoted $\mathbb{R}$ (respectively $\mathbb{R}_+$). We will also denote $\mathbb{R}_+ \cup \{0, \infty\}$ by $\mathbb{R}_\infty$.

A strict partial order on a set S is a transitive irreflexive relation on S. A pre-order on a set S is a transitive reflexive relation on S. If $\geq$ is a pre-order the relation $>$ on S defined by, for all $x, y \in S$, $x \geq y$ and $x \not\leq y$, is a strict partial order on S and the relation $\approx$ defined by $x \approx y$ if and only if $x \geq y$ and $x \leq y$ is an equivalence relation on S. The notation $a \gtrless b \Rightarrow c \gtrless d$ denotes if $a > b$ then $c > d$ and if $a < b$ then $c < d$. A strict partial order (or pre-order) $>$ on S in which any two distinct elements $s, t \in S$ are comparable, that is $s > t$ or $t > s$, is called a total order. A strict partial order $>$ on a set S is called well-founded if there are no infinite descending chains $s_1 > s_2 > s_3 > \ldots$ of elements of S. A strict partial order $\succ_E$ on S is said to be an extension of a strict partial order $\succ$ also on S, denoted by $\succ_E \supseteq \succ$, if for all $a, b \in S$, $a \succ b$ implies that $a \succ_E b$. An extension of a pre-order $\succeq$ on a set S is a pre-order $\succeq_E$ also on S, such that $\succ_E \supseteq \succ$ and for all $a, b \in S$ $a \approx b$ implies that $a \approx_E b$.

If $\{(S_i, >_i) | i \in \{1, \ldots, n\}\}$ is a sequence of strict partially ordered sets, then the combination $>= (>_1, >_2, \ldots, >_n)$ on the direct product $S_1 \times S_2 \times \cdots \times S_n$ is defined by $s = (s_1, s_2, \ldots, s_n) > t = (t_1, t_2, \ldots, t_n)$ if and only if there is an i with $s_i >_i t_i$, and $s_j = t_j$ for all $j < i$. Then $>$ is a strict partial order. If each $>_i$ is total, then $>$ is total. If each $>_i$ is well-founded, then $>$ is well-founded. The combination $>= (>_1, >_2, \ldots)$ on the direct product $S_1 \times S_2 \times \cdots$ of an infinite sequence of strict partially ordered sets is defined similarly; however, $>$ need not be well-founded even if each component is, as for example if each $S_i = \mathbb{N}$, we have $(1, 0, 0, \ldots) > (0, 1, 0, \ldots) > (0, 0, 1, \ldots) > \cdots$. Let $>$ be an order on T, let n be a fixed natural number, and let T^n be the set of n-tuples over T. Then the *lexicographic lifting from the left* of $>$ to T^n is the combination $>^{\mathrm{LexL}}= (>_1, >_2, \ldots)$ on $S_1 \times S_2 \times \cdots$ where $S_i = T$ and $>_i=>$, for $i = 1, \ldots, n$. The lexicographic combination $\geq_1; >_2$ of a pre-order $\geq_1$ and a pre-order or strict partial order $>_2$ on a set S (where $>_1$ is the associated strict partial order and $\approx_1$ the associated equivalence relation) is defined by, for $s, t \in S$, $s \geq_1; >_2 t$ if and only if either $s >_1 t$, or ($s \approx_1 t$ and $s >_2 t$).

If $\succ$ is a strict partial order on K and $\mu : S \longrightarrow K$ is a function, then the relation $>_\mu$ defined on S by $s >_\mu t$ if and only if $\mu(s) \succ \mu(t)$ is a strict partial order on S, called the strict partial order induced by μ, or the lifting of μ. This construction is often used when K is the real numbers, $\succ$ is the usual ordering on the reals and μ is some kind of weight function. If the restriction of $\succ$ to the image of μ is well-founded, then so is $>_\mu$. The relation $\geq_\mu$ defined on S by $s \geq_\mu t$ if and only if $\mu(s) \succ \mu(t)$ or $\mu(s) = \mu(t)$ is a pre-order on S, called the pre-order induced by μ. If $>$ is any order on S which is an extension of $>_\mu$, then $>=\geq_\mu; >$. Two ordered sets are said to be *order-isomorphic* if there is an order-preserving bijection between them.

2.1.1. TERMS

We work throughout with a finite set of function symbols $\mathcal{F}$ and a countable set of variables $\mathcal{X}$. We will use notions of both *fixed* and *variable arity* terms. In fixed arity terms, each $f \in \mathcal{F}$ is associated with a fixed natural number, called its *arity*. Symbols $f \in \mathcal{F}$ with arity 0 are called *constants*, which we will denote throughout by $\mathcal{F}_C \subseteq \mathcal{F}$. Symbols $g \in \mathcal{F}$ with arity 1 are called *unary*. A *fixed arity term* is either a constant, a variable, or an expression $f(t_1, \ldots, t_n)$ where f has arity n and $t_1, \ldots, t_n$ are terms. We denote the set of all fixed arity terms constructed from symbols in $\mathcal{F}$ and $\mathcal{X}$ by $T(\mathcal{F} \cup \mathcal{X})$. A term is called *monadic* if all of its function symbols are unary. We will often denote a monadic term $u(x) = a_1(a_2(\ldots(a_n(x))\ldots))$ by $a_1a_2\ldots a_n(x)$, and use the notation $u^m(x)$ to denote repeated application $u(\ldots(u(x))$. A *variable arity term* is either a variable, a function symbol or an expression $f(t_1, \ldots, t_n)$ where $f \in \mathcal{F}$ and $t_1, \ldots, t_n$ are terms. We denote the set of all variable arity terms by $V(\mathcal{F} \cup \mathcal{X})$.

We shall be investigating orders on fixed arity terms. We shall generally abuse notation and call fixed arity terms *terms*, and will only differentiate between fixed and variable arity terms when necessary. Let t be either a fixed or a variable arity term. We define the *head* of t (denoted $\mathrm{hd}(t)$) to be t if $t \in \mathcal{F} \cup \mathcal{X}$, or f if $t = f(t_1, \ldots, t_n)$. We define the *head arity* of t (denoted $\mathrm{hdar}(t)$) to be 0 if $t \in \mathcal{F} \cup \mathcal{X}$, or n if $t = f(t_1, \ldots, t_n)$. Note that if t is a fixed arity term, then $\mathrm{hdar}(t)$ is just the arity of $\mathrm{hd}(t)$. A term s is said to be a *subterm* of a term t if $s = t$ or $t = f(t_1, \ldots, t_n)$ and s is a subterm of one of the t_i. We call a subterm s of a term $t = f(t_1, \ldots, t_n)$ a *principal subterm of* t if $s = t_i$ for some $1 \leq i \leq n$.

Let t be a term. We define the *height of* t (denoted by $\mathrm{height}(t)$) to be 1 if $t \in \mathcal{F} \cup \mathcal{X}$ or $1 + \max\{\mathrm{height}(t_i) | 1 \leq i \leq n\}$ if $t = f(t_1, \ldots, t_n)$. A term in which no variable occurs is called a *ground term*. We denote the set of all ground fixed arity terms by $T(\mathcal{F})$, and the set of all ground variable arity terms by $V(\mathcal{F})$. Let $>$ be a strict partial order or pre-order on $T(\mathcal{F} \cup \mathcal{X})$. $>$ is called *monotonic* if $s > t$ implies that $f(s_1, \ldots, s, \ldots, s_n) > f(s_1, \ldots, t, \ldots, s_n)$ for all $s, t, s_i \in T(\mathcal{F} \cup \mathcal{X})$ and $f \in \mathcal{F}$.

A *substitution* $\sigma : \mathcal{X} \to T(\mathcal{F} \cup \mathcal{X})$ is a mapping from variables to terms. The image of x under σ is denoted $x\sigma$. The mapping σ on variables can be extended to a mapping on terms uniquely: $(f(t_1, \ldots, t_n))\sigma = f(t_1\sigma, \ldots, t_n\sigma)$ for all terms $f(t_1, \ldots, t_n)$. Let $\succ$ be a strict partial order or pre-order on $T(\mathcal{F} \cup \mathcal{X})$, and $s, t \in T(\mathcal{F} \cup \mathcal{X})$. We call $\succ$ *σ-stable* if $s \succ t$ implies $s\sigma \succ t\sigma$ for all ground substitutions $\sigma : \{x\} \to T(\mathcal{F})$, and stable if this property holds for all substitutions σ. A monotonic strict partial order or pre-order $>$ on $T(\mathcal{F} \cup \mathcal{X})$ is called a simplification order or simplification pre-order if it possesses the subterm property; i.e.

$$s = f(s_1, \ldots, s_i, \ldots, s_n) > s_i$$

for all terms s and $i = 1, \ldots, n$. A σ-stable simplification order is well-founded (see Dershowitz, 1982).

Let A be a set. We denote the set of all finite multisets with elements taken from A by $\mathrm{Mult}(A)$. Let $\succ$ be a strict partial order or pre-order on $T(\mathcal{F} \cup \mathcal{X})$, $\sim$ be an equivalence relation on $T(\mathcal{F} \cup \mathcal{X})$ with $\succ \cap \sim = \emptyset$, and $A, B \in \mathrm{Mult}(T(\mathcal{F} \cup \mathcal{X}))$. Then we define the *multiset extension* (denoted $\sqsupset$) of $\succ$ to be:

$$A \sqsupset B \Leftrightarrow A \setminus_M B \neq \emptyset \text{ and}$$
$$\forall t \in (B \setminus_M A)\ \exists s \in (A \setminus_M B) \text{ such that } s \succ t$$

where $A \setminus_M B$ is the multiset difference of A and B with respect to $\sim$, i.e. $A \setminus_M B = A$ if $\nexists x \in A, y \in B, x \sim y$ and $A \setminus_M B = (A \setminus \{x\}) \setminus_M (B \setminus \{y\})$ if $x \in A, y \in B$ and $x \sim y$.

2.1.2. THE HOMEOMORPHIC EMBEDDING RELATION

In this section we define the homeomorphic embedding relation for variable arity terms (Gallier, 1991), and $\#(u, v)$, the number of times the term u embeds in the term v.

DEFINITION 2.1. We define $\succeq_{\text{emb}}$, the *varyadic homeomorphic embedding relation*, on $V(\mathcal{F} \cup \mathcal{X})$ as follows: $t = g(t_1, \ldots, t_n) \succeq_{\text{emb}} f(s_1, \ldots, s_m) = s$ if and only if either:

(1) there is an i, $1 \leq i \leq n$ such that $t_i \succeq_{\text{emb}} s$, or
(2) $f = g$ and for all i, there exist j_i such that $t_{j_i} \succeq_{\text{emb}} s_i$ and $1 \leq j_1 \leq \cdots \leq j_m \leq n$.

We let $\succ_{\text{emb}}$ denote the strict part of $\succeq_{\text{emb}}$.

Note that the restriction of this definition to fixed arity terms is the usual homeomorphic embedding for fixed arity terms. We have (Dershowitz, 1982):

THEOREM 2.2. *Let $\mathcal{F}$ be a finite set of function symbols, $\mathcal{X}$ be a set of variables, $T = T(\mathcal{F} \cup \mathcal{X})$ be a term algebra and $\succeq_{\text{emb}}$ be the homeomorphic embedding relation on T. Then any simplification order $\succ$ on T is an extension of $\succ_{\text{emb}}$, and a strict partial order $>$ on T is well-founded if it extends $\succ_{\text{emb}}$.*

We introduce the following notation for conciseness.

NOTATION 2.3. Let $l, m, n, p, j_k \in \mathbb{N}$ for $1 \leq k \leq m$, such that $p \leq m \leq n$ and $l < n$. We will denote $j_p, \ldots, j_m$ is a strictly ascending subsequence of $l, l+1, l+2, \ldots, n$, i.e. $l \leq j_p < j_{p+1} < \cdots < j_m \leq n$, by $(j_k)_{k=p}^m \sqsubset [l, n]$.

DEFINITION 2.4. Let $u, v \in V(\mathcal{F})$. We define $\#(u, v)$, the number of times that u embeds in v, as follows:

Case 1 if $v \in \mathcal{F}$ then $\#(u, v) = 1$ if $u = v$ and 0 otherwise
Case 2 if $v = f(v_1, \ldots, v_n), u \in \mathcal{F}$ then $\#(u, v) = \#(u, f) + \sum_{k=1}^n \#(u, v_k)$
Case 3 if $v = f(v_1, \ldots, v_n), u = g(u_1, \ldots, u_m)$ then
Case 3a if $f \neq g$ or ($f = g$ and $m > n$) then
$\#(u, v) = \sum_{k=1}^n \#(u, v_k)$
Case 3b if $f = g$ and $n \geq m$ then
$\#(u, v) = \sum_{(j_k)_{k=1}^m \sqsubset [1,n]} (\prod_{i=1}^m \#(u_i, v_{j_i})) + \sum_{k=1}^n \#(u, v_k)$.

For example we have $\#(f(a,a), f(f(a), f(a))) = 1$, $\#(f(a,a), f(f(a,a))) = 1$ and $\#(f(a), f(f(a,a))) = 4$. We note that if $u, v \in T(\mathcal{F})$, then $\#(u,v)$ defines the number of times u homeomorphically embeds in v, since the restriction of varyadic embedding to $T(\mathcal{F})$ is the usual homeomorphic embedding relation for fixed arity terms.

2.2. ORDERS

We review definitions of standard orders.

2.2.1. POLYNOMIAL ORDERS

We follow definitions of polynomial interpretations and orders from Steinbach (1995) and Cropper and Martin (2000).

Suppose $a \in \mathbb{R}_+$ and $e_i \in \mathbb{N}$ for $1 \leq i \leq n$. Then $X_1^{e_1} X_2^{e_2} \ldots X_n^{e_n}$ is called a *monomial* over a set of *variables* $\{X_1, \ldots, X_n\}$. The degree of the monomial is the sum of the e_i. If M is a finite set of monomials, then a sum $p = \sum\{r * m \mid 0 \neq r \in \mathbb{R}, m \in M\}$ is called a *polynomial* over $\mathbb{R}$ on the variables $\mathcal{V} = \{X_1, \ldots, X_n\}$, written $p \in \mathbb{R}[X_1, \ldots, X_n]$. If $M = \{\ \}$, then p is represented by 0. A monomial occurs in the polynomial p if $m \in M$: a variable X occurs in p if X occurs in a monomial which occurs in p. A polynomial is called univariate if only one variable occurs in it. The degree $\deg(h)$ of a non-zero polynomial h is the degree of the largest monomial occurring in it.

A *polynomial interpretation* $[.] : \mathcal{F} \cup \mathcal{X} \to \mathbb{R}_+[\mathcal{V}]$ over $\mathbb{R}_+$ is defined by:

(1) assigning a polynomial to each n-ary function symbol f such that $[f](X_1, \ldots, X_n) \in \mathbb{R}_+[X_1, \ldots, X_n]$, where each X_i appears in at least one monomial with non-zero coefficient and
(2) $[x]() = X \in \mathcal{V}$ if $x \in \mathcal{X}$.

The polynomial interpretation $[.]$ can be extended to $[.] : T(\mathcal{F} \cup \mathcal{X}) \to \mathbb{R}_+[\mathcal{V}]$ by defining

$$[f(t_1, \ldots, t_n)] = [f]([t_1], [t_2], \ldots, [t_n]).$$

Let $\mathcal{F}_C \neq \emptyset$ and $[.]$ be a polynomial interpretation of $T(\mathcal{F} \cup \mathcal{X})$ over $\mathbb{R}_+$ and $\mu = \min\{[c] | c \in \mathcal{F}_C\}$. Then the order $>_{\text{poly}}$ is defined by, for $t, u \in T(\mathcal{F} \cup \mathcal{X})$, $t >_{\text{poly}} u$ if $[t](a_1, \ldots, a_n) > [u](a_1, \ldots, a_n)$ for all $a_1, \ldots, a_n \in \mathbb{R}_+$ such that $a_1, \ldots, a_n \geq \mu$. There are a variety of additional conditions which may be imposed upon $[.]$ to ensure that $>_{\text{poly}}$ is σ-stable or a simplification order: see Steinbach (1995).

If $\mathcal{F}$ consists of unary function symbols, then we may define a different order on T, the eventually dominates order $>_{\text{Dpoly}}$ (Cropper and Martin, 2000). Let $A(x) = \sum_{i=0}^{n} a_i X^i$, $B(x) = \sum_{i=0}^{m} b_i X^i$ with $a_i, b_i \in \mathbb{R}$, $n, m \geq 1$ and $a_n, b_m \neq 0$. Then $A(x)$ is said to eventually dominate $B(x)$ if there is a $\chi(A, B) \in \mathbb{R}$ such that $A(r) > B(r)$ for all $r > \chi(A, B)$, or equivalently if either $n > m$ or $n = m$ and $(a_n, \ldots, a_0)$ is greater than $(b_m, \ldots, b_0)$ in the lexicographic order from the left. This order is total on univariate real polynomials.

DEFINITION 2.5. Let $T = T(\mathcal{F} \cup \{x\})$, and let $[.]$ be a polynomial interpretation of $\mathcal{F}$ over $\mathbb{R}_+$. Then the order $>_{\text{Dpoly}}$ on T is defined by, for all $s, t \in T$, $s >_{\text{Dpoly}} t$ if and only if $[s](X)$ eventually dominates $[t](X)$.

In this case it is somewhat easier to analyse when $>_{\mathrm{Dpoly}}$ is a σ-stable simplification order. We have (Cropper and Martin, 2000):

THEOREM 2.6. *Let $\mathcal{F}$, $[.]$ and $>_{\mathrm{Dpoly}}$ be as above and suppose that for each $f \in \mathcal{F}$ either $\deg([f(x)]) > 1$ or $[f(x)] = a_f X + b_f$ with $a_f > 1$ or $a_f = 1$ and $b_f > 0$. Then $>_{\mathrm{Dpoly}}$ is a simplification order which is stable under substitution, and if further $[.]$ is injective on T then $>_{\mathrm{Dpoly}}$ is total on T.*

2.2.2. Knuth–Bendix order

In this section we define the Knuth–Bendix order, which first appeared in Knuth and Bendix (1970). This order has been much studied: see, for instance, Dick *et al.* (1990) and Steinbach (1995).

DEFINITION 2.7. Let ρ be a positive real number and let $\mu : \mathcal{F} \cup \mathcal{X} \to \mathbb{R}_+ \cup \{0\}$ satisfy:

(1) for all variables x, $\mu(x) = \rho$,
(2) for all $c \in \mathcal{F}_c$, $\mu(c) \geq \rho$, and
(3) at most one unary function symbol $f \in \mathcal{F}$ has $\mu(f) = 0$.

We extend μ to terms $T(\mathcal{F} \cup \mathcal{X})$ by the homomorphism

$$\mu(f(t_1, \ldots, t_n)) = \mu(f) + \sum_{i=1}^{n} \mu(t_i).$$

Let $\succ$ be a precedence on $\mathcal{F}$ such that if there is a unary $f \in \mathcal{F}$ with weight 0, then $f \succ g$ for each $g \in \mathcal{F} \setminus \{f\}$. Then the *Knuth–Bendix order* $>_{\mathrm{kbo}}$ is defined by, for $t, u \in T(\mathcal{F} \cup \mathcal{X})$: $t >_{\mathrm{kbo}} u$ if for all $x \in \mathcal{X} : \#(x, t) \geq \#(x, u)$ and either:

(1) $\mu(t) > \mu(u)$ or
(2) $\mu(t) = \mu(u)$ and
   (a) $t = f^n(x)$ and $u = x$ for some $n \geq 1$ and $x \in \mathcal{X}$, or
   (b) $t = h(t_1, \ldots, t_n)$, $u = g(u_1, \ldots, u_m)$ and $h \succ g$, or
   (c) $t = h(t_1, \ldots, t_n)$, $u = h(u_1, \ldots, u_n)$ and $(t_1, \ldots, t_n) >_{\mathrm{kbo}}^{\mathrm{LexL}} (u_1, \ldots, u_n)$.

The following proposition follows from Knuth and Bendix (1970).

PROPOSITION 2.8. *The Knuth–Bendix order is a σ-stable simplification order.*

2.2.3. the recursive path order

In this section we define the recursive path order (Dershowitz, 1982). Let $\mathcal{F}$ be a finite set of function symbols equipped with a pre-order $\succ$ and $\mathcal{X}$ be a finite set of variables. We define $\geq_{\mathrm{rpo}}$, the *recursive path order* on $T(\mathcal{F} \cup \mathcal{X})$ as follows. Let $s = f(s_1, \ldots, s_n)$, $t = g(t_1, \ldots, t_m)$ where $n = \alpha(f) \geq 0$ and $m = \alpha(g) \geq 0$. Then $s \geq_{\mathrm{rpo}} t$ if and only if one of the following holds

(1) $f \succeq g$ and $s >_{\mathrm{rpo}} t_i$ for $i = 1, \ldots, m$

(2) $f \approx g$ and $\{\!\{s_1, \ldots, s_n\}\!\} \sqsupseteq_{\mathrm{rpo}} \{\!\{t_1, \ldots, t_m\}\!\}$
(3) $\neg(f \succeq g)$ and there is a j such that $s_j \geq_{\mathrm{rpo}} t$.

where $\sqsupseteq_{\mathrm{rpo}}$ is the multiset extension of $\geq_{\mathrm{rpo}}$. We denote by $\geq_{\mathrm{rpo}}$ the pre-order thus defined, and $>_{\mathrm{rpo}}$ the related strict partial order, so $s >_{\mathrm{rpo}} t$ if and only if $s \geq_{\mathrm{rpo}} t$ and $\neg(t \geq_{\mathrm{rpo}} s)$.

We have (Dershowitz, 1982):

Proposition 2.9. *Let $\geq_{\mathrm{rpo}}$ be defined on $T = T(\mathcal{F} \cup \mathcal{X})$ as above.*

(1) $\geq_{\mathrm{rpo}}$ is a simplification pre-order on T.
(2) if $s = f(s_1, \ldots, s_n), t = f(t_1, \ldots, t_n)$ and there is a j with $1 \leq j \leq n$ such that $s_j >_{\mathrm{rpo}} t_i$ for $1 \leq i \leq n$, then $s >_{\mathrm{rpo}} t$.

Note that, for example, $f(f(a, a), a) \approx f(a, f(a, a))$ so $>_{\mathrm{rpo}}$ is not total.

2.2.4. orders on vector spaces

In what follows we shall need orders on $\mathbb{R}^n$: we recall some background from Martin (1989, 1995).

A strict partial order on $\mathbb{R}^n$ is called *monotonic* if $u > v$ implies $u + w > v + w$ for all $u, v, w \in \mathbb{R}^n$, and a *division order* if further each $E_p > \underline{0}$ for $1 \leq p \leq n$, where E_p is the pth coordinate vector, that is the vector with a 1 in the pth position and zeros elsewhere, and $\underline{0}$ is the zero vector. The restriction to $\mathbb{N}^n$ of any division order is well-founded.

To obtain strict partial orders on $\mathbb{R}^n$ we proceed as follows (see Martin, 1989). Let A be any real matrix with n rows. The strict partial order $>_A$ is defined by $u >_A v$ if and only if uA is greater than vA in the lexicographic order from the left. It is a monotonic order. If further each coordinate vector is greater than $\underline{0}$ it is a division order. All total monotonic and division orders arise in this way, and each orthogonal matrix A gives rise to a distinct such strict partial order. The $n!$ orthogonal matrices obtained by permuting the rows of the identity matrix give $n!$ lexicographic orders on $\mathbb{R}^n$, which are just the lexicographic extensions of the $n!$ possible permutations of the standard basis vectors.

Vectors are sometimes "ordered by weight", that is a fixed vector h of non-negative real weights is given and the strict partial order is defined by $u > v$ if and only if $u.h > v.h$, where . is the usual dot product of vectors. Orders by weight are equivalent up to scalar multiples of w. An orthogonal matrix whose first column is the corresponding vector of weights (suitably normalized) gives rise to a total monotonic or division order which extends such an order by weight. Conversely given an orthogonal matrix as above, its first column defines a weight order which the matrix order extends.

Multiplication by a non-negative scalar defines an equivalence relation on non-zero n-vectors over the non-negative reals. The set of equivalence classes bijects with a topological object, the non-negative orthant in projective n-space, which we denote by $\mathcal{P}_n$. This analysis of the weights means that we have a surjective mapping from the set of monotonic orderings of $\mathbb{R}^n$ to $\mathcal{P}_n$: thus $\mathcal{P}_n$ forms a classifying space for this set of orderings.

Thus, given a set of pairs of vectors (u_i, v_i), we may decide whether or not there is an order by weight which orders each pair from left to right by deciding if there is a vector w with non-negative entries and $u_i.w > v_i.w$ for each i: this is a standard linear programming technique. If there is no such w, then it will not be possible to find such an order. In terms of the classifying space, this is equivalent to determining whether or not a certain subset is empty.

In the case $n = 1$ our theory reduces to saying that there is one division order on $\mathbb{R}$: the usual linear order denoted by $>_l$. There are two further monotonic orders: the trivial order and the reverse of $>_l$.

2.3. ORDERS ON MONADIC TERMS

In this section we will review some recent work on classifying orders on monadic terms over a finite set $A = \{a_1, \ldots, a_n\}$ of unary function symbols. The results will be applied in Section 3 to prove our results about orders on terms. The first two theorems are from Scott (1994).

THEOREM 2.10. *Let $A = \{a_1, \ldots, a_n\}$ be a set of unary function symbols, x be a variable and let $\succ$ be a total stable simplification order on $T = T(A \cup \{x\})$. Given $u(x), v(x) \in T \setminus \{x\}$ there exists τ such that $0 \le \tau \le \infty$ and for all $i, j \ge 1$ we have*

$$u^i(x) \gtrless v^j(x) \qquad \text{if } i \gtrless j\tau.$$

THEOREM 2.11. *Let $A = \{a_1, \ldots, a_n\}$ be a set of unary function symbols, x be a variable, and let $\succ$ be a total stable simplification order on $T = T(A \cup \{x\})$, so without loss of generality we may assume that $a_1(x) \prec a_2(x) \prec \cdots \prec a_n(x)$. Then there exist $\mu_1, \ldots, \mu_n$ with $0 \le \mu_1 \le \cdots \le \mu_n = 1$, such that if $\mu : T \to \mathbb{R}$ is defined by*

$$\mu(u) = \sum_{i=1}^{n} \mu_i \, \#(a_i, u),$$

then $\succ$ is an extension of the order $>_\mu$ induced on T by lifting μ.

The order $\succ$ on T is said to be linearizable, the sequence $(\mu_1, \ldots, \mu_n)$ is called a weight sequence for $(T, \succ)$ and the function $\mu : T \to \mathbb{R}$ is called a weight function for $\succ$, and defines a weight pre-order $\succeq_\mu$ on T such that $\succ = \succeq_\mu ; \succ$. To compute the μ_i note that by Theorem 2.10 for each $r = 1, \ldots, n-1$, there is a τ_r with $0 \le \tau_r \le 1$ such that $a_{r+1}^i(x) \gtrless a_r^j(x)$ if $i \gtrless j\tau_r$. Let $\mu_n = 1$ and $\mu_r = \tau_{n-1} \cdots \tau_r$ for $i = 1, \ldots, n-1$. We note that $\mathcal{P}_n$ forms a classifying space for the set of total stable simplification orderings on T: the reasoning is as for vector spaces in Subsection 2.2.4.

In the case of two letters, Martin and Scott (1997) refines Theorem 2.11 somewhat as follows: we use this result in developing our examples in Section 4.

DEFINITION 2.12. Let $\mathcal{F} = \{f, g\}$ be a set of unary function symbols. Given a real number $\lambda > 0$, the pre-order $\trianglerighteq_\lambda$ on $T(\mathcal{F} \cup \{x\})$ with respect to $\{f, g\}$ is defined by

$$g^{p_0} f g^{p_1} f \ldots f g^{p_m}(x) \trianglerighteq_\lambda g^{q_0} f g^{q_1} f \ldots f g^{q_n}(x)$$

if and only if either $m > n$, or $m = n$ and

$$p_n \lambda^n + \cdots + p_1 \lambda + p_0 \geqslant q_n \lambda^n + \cdots + q_1 \lambda + q_0.$$

We abuse notation and extend this to the case $\lambda = 0, \infty$ in the obvious way.

Thus, a λ pre-order first orders by the number of f's and then by the number of g's biased according to position among the f's, so that in particular if $i > j\lambda$ then $g^i f(x) \trianglerighteq_\lambda fg^j(x)$ and if $i < j\lambda$ then $fg^j(x) \trianglerighteq_\lambda g^i f(x)$.

Theorem 2.13. *Let $\succ$ be a total stable simplification order on $T(\mathcal{F} \cup \{x\})$ and suppose that $f(x) \succ g(x)$. Then there is a constant μ_1 with $1 \geq \mu_1 \geq 0$ such that $\succ = \geq_\mu ; \succ$ where $(1, \mu_1)$ is a weight sequence for $\succ$. Then either $\mu_1 > 0$, or $\mu_1 = 0$ in which case there is a real constant $\infty \geq \lambda \geq 0$ such that $\succ = \geq_{\mu_1} ; \trianglerighteq_\lambda ; \succ$.*

3. Linearizable Orders

In this section we prove our main theorem concerning the linearization of term orders. We proceed by defining the notion of an independent set and a one-parameter family over a term algebra T. We then show that if T is equipped with a σ-stable simplification order satisfying certain conditions, then any one-parameter family is order-isomorphic as an algebra to a certain algebra on unary function symbols equipped with a σ-stable simplification order. This means that we may apply the results for strings explained above to the latter, and hence read off our results. We then address the question of lifting orders on the spine to orders on the whole term algebra.

3.1. the main theorem

We now establish the basic definitions for the main result. Our first task is to define a one-parameter family: this will be a subalgebra which is isomorphic to a term algebra on monadic terms. Thus we need candidates for the generators of the one-parameter family: these are the one-parameter terms. We need conditions on the subalgebra they generate to ensure that we obtain a free algebra: this gives us the notion of an independent set.

Let $T = T(\mathcal{F} \cup \mathcal{X})$ be a term algebra with function symbols $\mathcal{F}$ and variables $\mathcal{X}$. An element of $T \setminus \mathcal{X}$ is called a *one-parameter term* over T if it contains exactly one variable x of $\mathcal{X}$, which may occur more than once. If $\mathcal{G} = \{{}_jG(x) | j \in I\}$ is a set of one-parameter terms and $x \in \mathcal{X}$ let *the span of* $\mathcal{G}$ be:

$$T_{\mathcal{G},x} = \{x\} \cup \{{}_{j_1}G^{i_1}({}_{j_2}G^{i_2}(\cdots({}_{j_k}G^{i_k}(x))\cdots) | 1 \leq m \leq k, j_m \in I, i_m \in \mathbb{N}\}.$$

We call $\mathcal{G}$ *independent* if for all $s, t \in T_{\mathcal{G},x}$ with

$$s = {}_{j_1}G^{i_1}({}_{j_2}G^{i_2}(\cdots({}_{j_k}G^{i_k}(x))\cdots)), \qquad t = {}_{j'_1}G^{i'_1}({}_{j'_2}G^{i'_2}(\cdots({}_{j'_{k'}}G^{i'_{k'}}(x))\cdots)),$$

then $s = t$ if and only if $k = k'$ and for all p, $1 \leq p \leq k$, we have $i_p = i_{p'}$ and $j_p = j_{p'}$. We call $S \subseteq T$ a *one-parameter family in* T if $S = T_{\mathcal{G},x}$ for some $x \in \mathcal{X}$ and $\mathcal{G} \subseteq T$ is an independent set.

As an example for $f \in \mathcal{F} \setminus \mathcal{F}_C$ let $\hat{f}(x) = f(x, x, \ldots, x)$. Let $\mathcal{H} = \{\hat{f}(x) | f \in \mathcal{F} \setminus \mathcal{F}_C\}$. Then $T_{\mathcal{H},x}$ is a one-parameter family, called the *spine* of T. So in the running example of the introduction the one-parameter terms are $\{F(x) = f(x,x), G(x) = g(x,x)\}$ and they form an independent set and generate a one-parameter family.

Now suppose that $\mathcal{G} = \{{}_jG | j \in I\}$. We define the *associated unary algebra* S_A as $S_A = T(A \cup \{y\})$, where $A = \{a_i | i \in I\}$ is a set of unary function symbols and y is a variable. We define $\alpha : S \to S_A$ by $\alpha : x \to y$, and $\alpha : {}_iG(z) \to a_i(\alpha(z))$ for all $i \in I, z \in S$. We have the following.

Lemma 3.1. *Given $T, \mathcal{G}, S, S_A, \alpha$ as above, then $\alpha : S \to S_A$ is a bijection.*

Proof. Routine. □

Now suppose that $>$ is an order on T. Then $>$ defines a relation $>_\alpha$ on S_A by

$$s >_\alpha t \Leftrightarrow s = \alpha(s'), t = \alpha(t'), s', t' \in S \text{ and } s' > t'.$$

We have the following.

Theorem 3.2. *Let $T, \mathcal{G}, S, S_A, \alpha$ be as above, and suppose $>$ is a strict partial order on T. Then:*

(1) $>_\alpha$ is a strict partial order on S_A
(2) $>_\alpha$ is total if and only if $> |_S$ is total
(3) $>_\alpha$ is a simplification order if and only if $> |_S$ is a simplification order
(4) $>_\alpha$ is stable under substitution if and only if $>$ is preserved under σ for any $\sigma : \{x\} \to S$.

Proof. Routine. □

Thus $(S, > |_S)$ is order isomorphic to $(S_A, >_\alpha)$, and abusing notation we may thus regard S as both $T_{\mathcal{G},x}$ and as a subset of $T = T(\mathcal{F} \cup \mathcal{X})$. In particular for $s \in S$ we may define both $\#(f, s)$ for $f \in \mathcal{F}$ and $\#(G, s)$ for $G \in \mathcal{G}$.

Applying the results on monadic terms quoted in Section 2 to S_A, and lifting to S via α using Theorem 3.2, we now obtain analogues of these results for terms.

Theorem 3.3. *Let $T, \mathcal{G}, S$ be as above and let $\succ$ be a σ-stable simplification order on T which is total on restriction to S. Let $u(x), v(x)$ be elements of S. Then there is a $\tau \in \mathbb{R}_\infty$ such that for all $p \in T$, $u^i(p) \gtrless v^j(p)$ if $i \gtrless j\tau$.*

Proof. The result holds for all $p \in S$ via the lifting and for all $p \in T$ as $\succ$ is σ-stable. □

In our running example if we take $\succ$ to be the polynomial order induced by $[f(x, y)] = 2XY^2, [g(x, y)] = X^2 + Y^2, [d] = D, [x] = X, [y] = Y$ and $u(x) = F(x), v(x) = G(x)$ we have $\tau = \log 2/\log 3$.

The next two results follow from Theorems 2.10 and 2.11.

Theorem 3.4. *Let T, $\mathcal{G}$, S and $\succ$ be as above. Then $\succ |_S$ is linearizable over S, with classifying space $\mathcal{P}_n$.*

Proof. As $\succ$ is a σ-stable simplification order and $\succ |_S$ is total on S it corresponds to a total σ-stable simplification order $\succ_\alpha$ on S_A, for which the required μ_i exist by Theorem 2.11. Thus if $u(x), v(x)$ in S satisfy $u^i(x) \gtrless v^j(x)$ if $i \gtrless j\tau$, then $\alpha(u)$, $\alpha(v)$ satisfy the same equation with a_i for ${}_iG$, so $\alpha(u) \gtrless_\alpha \alpha(v)$ and it follows that $u \gtrless v$. □

As $\succ |_S$ is total on S, we may assume without loss of generality that $\mathcal{G} = \{{}_1G, \ldots, {}_nG\}$ where ${}_nG(x) \succ {}_{n-1}G(x) \succ \cdots \succ {}_1G(x)$, that $(S, \succ |_S)$ has a weight sequence satisfying $1 = \mu_n \geq \mu_{n-1} \geq \cdots \geq \mu_1 \geq 0$, and weight function μ on S given by

$\mu(s) = \sum_{{}_iG \in \mathcal{G}} \#({}_iG, s)\mu_i$, and that $\succ|_S = \succeq_\mu; \succ$. In our running example $1 = \mu_2 > \mu_1 = \log 2/\log 3 > 0$, so for example $\mu(G(F(x))) = 1 + \log 2/\log 3$.

We also have the following.

Theorem 3.5. *Let T, $\mathcal{G}$, S and $\succ$ be as above. Then there exist permutations π and ψ of $\{1, \dots, n\}$ such that, for $s \in S$,*

$$_{\pi(1)}G^{t_{\pi(1)}}\,{}_{\pi(2)}G^{t_{\pi(2)}} \dots {}_{\pi(n)}G^{t_{\pi(n)}}(x) \succeq s \succeq {}_{\psi(1)}G^{t_{\psi(1)}}\,{}_{\psi(2)}G^{t_{\psi(2)}} \dots {}_{\psi(n)}G^{t_{\psi(n)}}(x)$$

where $t_i = \#({}_iG, s)$.

Proof. The analogous result for strings was proved in Martin (1990): translating via α as in the previous theorem gives our result. □

Thus, in our running example we have that an element s of S with $\#(F, s) = n$, $\#(G, s) = m$ satisfies $F^n G^m \succeq s \succeq G^m F^n$.

3.2. the spine of $T(\mathcal{F} \cup \mathcal{X})$

In this section we show how stronger results can be obtained in the case where $S = T_{\mathcal{H},x}$, the spine of $T = T(\mathcal{F} \cup \mathcal{X})$. We first show that any total simplification order $\succ$ on S can be lifted to a simplification pre-order $\trianglerighteq$ on $T(\mathcal{F})$ which is compatible with $\succ$ on ground images of S, that is if $s, t \in S$ and $s \succ t$, then $s\sigma \trianglerighteq t\sigma$ for any substitution $\sigma : \{x\} \to T(\mathcal{F})$.

We already know that any order $\succ$ on T which is a total simplification order, stable under substitution on restriction to $S = T_{\mathcal{H},x}$, induces a weight function μ on S, so that $\succ|_S = \geq_\mu; \succ$. We show further that, under minor restrictions on $\mathcal{F}$ and $\succ$, μ can be lifted to a function $\bar{\mu} : T(\mathcal{F}) \to \mathbb{R}$ such that $\succ \subseteq \geq_{\bar{\mu}}$, so that if $u, v \in T(\mathcal{F})$ and $u \succ v$, then $\bar{\mu}(u) \geq \bar{\mu}(v)$.

To make these results precise takes some further technicalities, which arise because S contains no constants. We suppose that $T = T(\mathcal{F} \cup \mathcal{X})$ and $S = T_{\mathcal{H},x}$ as above, and suppose that $d \in \mathcal{F}_C$. Then S bijects with $S_d = T_{\mathcal{H},d} = \{s(d) | s(x) \in S\}$, and any order $\succ$ on S induces an order $\succ'$ on S_d by $s(d) \succ' t(d)$ if and only if $s(x) \succ t(x)$. Thus the bijection is an order isomorphism, so if $\succ$ is total or a simplification order, then $\succ'$ will also have this property.

Thus we have the following.

Theorem 3.6. *Let T, $\mathcal{H}$ be as above, and let $\succ$ be a total simplification order on $S = T_{\mathcal{H},x}$, preserved under substitutions $\sigma : \{x\} \to S$. Suppose $\mathcal{F}_C \neq \emptyset$. Then there is a simplification pre-order $\trianglerighteq$ on $T(\mathcal{F})$ such that if $s, t \in S$ and $s \succ t$, then $s\sigma \trianglerighteq t\sigma$ for any $\sigma : \{x\} \to T(\mathcal{F})$.*

Proof. Let $d \in \mathcal{F}_C$. Then S_d bijects with S and we may define $\succ$ on S_d by $u(d) \succ v(d)$ if and only if $u(x) \succ v(x)$. For each u in $T(\mathcal{F})$ we define $u' \in S_d$, and let $u \trianglerighteq v$ if and only if $u' \succ v'$ or $u' = v'$. To define u' we proceed recursively over the structure of u. So if $a \in \mathcal{F}_C$, let $a' = d \in S_d$. If $u = f(u_1, \dots, u_n) \in T$ and $\{u'_1, \dots, u'_n\}$ is the corresponding subset of S_d, then $\{u'_1, \dots, u'_n\}$ has a unique maximal element U, since S_d is totally ordered under $\succ$. Let $u' = f(U, \dots, U) \in S_d$.

It is clear that $\trianglerighteq$ is transitive and reflexive. If $u = f(u_1, \dots, u_n)$, then $u' = f(U, \dots, U)$

where $U = \max_{\succeq}\{u'_1, \ldots, u'_n\}$ so $u' \succ U \succeq u'_i$ for all i. So $u \trianglerighteq u_i$ for all i and $\trianglerighteq$ has the subterm property.

To show that $\trianglerighteq$ is monotonic suppose $u, v \in T(\mathcal{F})$ with $u \trianglerighteq v$, so $u' \succeq v'$. Consider $p = f(u, t_2, \ldots, t_n)$ and $q = f(v, t_2, \ldots, t_n)$, where $p' = f(U, \ldots, U)$, $q' = f(V, \ldots, V)$ and $U = \max_{\succeq}\{u', t'_2, \ldots, t'_n\}$, $V = \max_{\succeq}\{V', t'_2, \ldots, t'_n\}$. Let $t = \max_{\succeq}\{t'_2, \ldots, t'_n\}$. There are four cases to consider:

(1) **Case** $t \succ u'$ **and** $t \succ v'$. In this case $U = V = t$ so $p' = q'$, and it follows that $p' \succeq q'$ and $p \trianglerighteq q$.
(2) **Case** $t \preceq u'$ **and** $t \succ v'$. In this case $U = u' \succeq t = V$ so $p' \succeq q'$ and $p \trianglerighteq q$.
(3) **Case** $t \succ u'$ **and** $t \preceq v'$. This case cannot occur as $v' \succeq t \succ u' \succeq v'$, so $v' \succ v'$, which is impossible.
(4) **Case** $t \preceq u'$ **and** $t \preceq v'$. In this case $U = u'$, $V = v'$, so $U \succeq V$ and $p' \succeq q'$, so $p \trianglerighteq q$.

So in each case, $p \trianglerighteq q$ as required. Thus $\trianglerighteq$ is a simplification pre-order.

If $\sigma : \{x\} \to T(\mathcal{F})$, then $(x\sigma)' = q(d) \in S_d$ and if s, t are elements of S, then $[s(x\sigma)]' = s((x\sigma)') = s(q(d))$ and similarly $[t(x\sigma)]' = t(q(d))$. Now $s(x) \succ t(x)$, so $s(d) \succ t(d)$ and hence $s(q(d)) \succ t(q(d))$ because of the stability property of the hypothesis, so $s(q(d)) \succ t(q(d))$ and hence $s(x\sigma) \trianglerighteq t(x\sigma)$ as required. □

For our second result we proceed as follows.

Theorem 3.7. *Let $T, \mathcal{H}, S$ be as above, and $\succ$ an order on T which is a total simplification order on restriction to S, preserved under substitutions $\sigma : \{x\} \to T(\mathcal{F})$, with weight function μ on S. Suppose that $\mathcal{F}_C$ has a unique maximal element d. Then there is a function $\bar{\mu} : T(\mathcal{F}) \to \mathbb{R}$ such that:*

(1) $\bar{\mu}(s(d)) = \mu(s(d))$ for all $s(d) \in S_d$,
(2) $\succ \subseteq \geq_{\bar{\mu}}$ so that if $u \succ v$, then $u \geq_{\bar{\mu}} v$.

Proof. As before we work with S_d, which is order-isomorphic to S since $\succ$ is preserved under substitutions $\sigma : \{x\} \to T(\mathcal{F})$. We show first that every element of $T(\mathcal{F})$ has an upper bound in S_d, that is that there is an element $\bar{u} \in S_d$ with $u \preceq \bar{u}$. The proof is by induction on the height of u. If u is a constant, then $u \preceq d$, so set $\bar{u} = d$. If $u = f(u_1, \ldots, u_n)$, then by induction each u_i has an upper bound $\bar{u_i}$ in S_d, and as $\preceq$ is total on S_d, the set $\{\bar{u_1}, \ldots, \bar{u_n}\}$ has a unique maximal element U. Thus $u = f(u_1, \ldots, u_n) \preceq f(\bar{u_1}, \ldots, \bar{u_n}) \preceq f(U, \ldots, U) \in S_d$ as required.

Now if u is any element of $T(\mathcal{F})$ the upper bounds of u in S_d form a non-empty subset of S_d. Now $\prec$ is total on S_d, so this subset has a unique minimal element $\bar{u}$, the least upper bound of u in S_d: that is to say that $u \preceq \bar{u}$ and if $u \preceq h$ and $h \in S_d$, then $\bar{u} \preceq h$. Thus we can define $\bar{\mu}$ as follows. If $s(d) \in S_d$, then $\bar{\mu}(s(d)) = \mu(s(x)) \in \mathbb{R}$. If $u \in T(\mathcal{F})$, then $\bar{\mu}(u) = \bar{\mu}(\bar{u}) = \mu(s(x))$ where $\bar{u} = s(d) \in S_d$. Since $\succ$ is total on S, any two elements of T are comparable under $\geq_{\bar{\mu}}$.

Now suppose that $u \succ v$. If $u \geq_{\bar{\mu}} v$, then we have the required result, so we may suppose that $v >_{\bar{\mu}} u$, that is to say that $\bar{\mu}(v) > \bar{\mu}(u)$, that is $\bar{\mu}(\bar{v}) > \bar{\mu}(\bar{u})$ so $\mu(V(x)) > \mu(U(x))$ where $\bar{v} = V(d), \bar{u} = U(d)$. Hence $V(x) \succ U(x)$ so $\bar{v} \succ \bar{u}$. Thus as $\bar{u} \succeq u$ we have $\bar{v} \succ \bar{u} \succeq u \succ v$. Now consider $\bar{v}$. This is the least upper bound of v in S_d. But $\bar{u} \succ v$, $\bar{u}$

also lies in S_d and $\bar{v} \succ \bar{u}$, which contradicts the fact that $\bar{v}$ is the least upper bound of v. Thus $u \geq_{\bar{\mu}} v$ as required. □

Thus, in our running example $\bar{\mu}(g(g(d,d), f(d,d))) = \mu(G(F(x))) = 1 + \log 2/\log 3$.

4. One-parameter Families

In this section we note some examples of one-parameter families $T_{\mathcal{G},x}$ over term algebras $T = T(\mathcal{F} \cup \mathcal{X})$ which satisfy the hypotheses of our main results and will be used below. We then investigate the Knuth–Bendix and polynomial orders in the light of the results above.

4.1. examples

We have already seen that the spine of $T = T(\mathcal{F} \cup \mathcal{X})$ is a one-parameter family. In Examples 4.1–4.5 below it is straightforward to prove by induction that $T_{\mathcal{G},x}$ is a one-parameter family, and there exist σ-stable simplification orders on T which are total on $T_{\mathcal{G},x}$. Example 4.6 shows a one-parameter family which does not admit an order of the required kind.

Example 4.1. Let $x \in \mathcal{X}$ and let $\mathcal{G}$ be a non-empty subset of the spine of T,

$$\mathcal{G} \subseteq \{\phi_f(x) = f(x, \ldots, x) | f \in \mathcal{F} \setminus \mathcal{F}_C\}.$$

Note that $\phi_f^n(x)$ is the unique maximal tree under embedding of height $n+1$ in $T(\{f, x\})$.

Example 4.2. Let $x \in \mathcal{X}$ and let $\mathcal{G}$ be a non-empty subset of

$$\{\theta_f(x) = f(x, a, \ldots, a) | f \in \mathcal{F} \setminus \mathcal{F}_C\}.$$

Note that $\theta_f^n(x)$ is a minimal tree under embedding of height $n+1$ in $T(\{f, a, x\})$.

Example 4.3. This example subsumes the two previous examples. Let $x \in X$, $f \in \mathcal{F} \setminus \mathcal{F}_C$ and let $\hat{f}(x)$ be any term containing at least one occurrence of both f and x, so $\hat{f}(x) \in T(\{f, x\} \cup \mathcal{F}_C) \setminus (T(\{f\} \cup \mathcal{F}_C) \cup \{x\})$. Let $\mathcal{G}$ be a non-empty subset of

$$\{\hat{f}(x) | f \in \mathcal{F} \setminus \mathcal{F}_C\}.$$

Example 4.4. Let $f \in \mathcal{F} \setminus \mathcal{F}_c$, $x \in \mathcal{X}$ and suppose $\mathcal{F}_c \neq \emptyset$ and let $\mathcal{G}$ be a non-empty subset of

$$\{f(a, x) | a \in \mathcal{F}_c\}.$$

Example 4.5. More generally, let $f \in \mathcal{F} \setminus \mathcal{F}_c, x, y \in \mathcal{X}$ and suppose $\mathcal{F}_c \neq \emptyset$. Let $F(x, y) \in T(\{f, x, y\})$ and let $\mathcal{G}$ be a non-empty subset of

$$\{F(x, a) | a \in \mathcal{F}_c\}.$$

Example 4.6. Each of the examples above has some kind of constraint to ensure independence and σ-stability. The latter will fail if two distinct terms in $T_{\mathcal{G},x}$ are unifiable, that is become equal under substitution. For example consider $T = T(\{f, a, x\})$, and let $\mathcal{G} = \{F(x) = f(a, x), G(x) = f(x, a)\}$. Then $T_{\mathcal{G},x}$ is a one-parameter family, but in any total order $\succ$ on T, we have either $F(x) \succ G(x)$ or $G(x) \succ F(x)$, but under the substitution $\sigma : x \to a$ we have $\sigma(F(x)) = \sigma(G(x))$, violating σ-stability.

4.2. Knuth–Bendix order

We now investigate the Knuth–Bendix order in this framework. We have shown in Theorem 3.4 that any σ-stable simplification order over a one-parameter family gives rise to a weight sequence. We now show that given any sequence $0 < \mu_1 \leq \mu_2 \leq \cdots \leq \mu_n = 1$ then it arises as the weight sequence of the restriction of a Knuth–Bendix order to a one-parameter family $S = T_{\mathcal{G},x} \subset T(\mathcal{F} \cup \mathcal{X})$ where $|\mathcal{G}| = n$.

We then investigate Knuth–Bendix orders in more detail in the case that $|\mathcal{G}| = 2$, extracting for any such $\mathcal{G}$ and any Knuth–Bendix order the τ, λ of Theorem 2.13 and the weight sequence of Theorem 2.11.

To prove our first result we need to construct a Knuth–Bendix order.

Theorem 4.7. *Let $0 < \mu_1 \leq \cdots \leq \mu_n = 1$ and $T = T(\mathcal{F} \cup \mathcal{X})$ where $|\mathcal{F} \setminus \mathcal{F}_C| = n$ and $|\mathcal{F}_C| \geq 1$. Then there is a $\mathcal{G} \subset T$ with $|\mathcal{G}| = n$, a one-parameter family $S = T_{\mathcal{G},x}$ and a simplification order $>$ on T which is total on restriction to S and has weight sequence $(\mu_1, \ldots, \mu_n)$ over S.*

Proof. Let $a \in \mathcal{F}_C$. For each $f_i \in \mathcal{F}\backslash\mathcal{F}_C = \{f_1, \ldots, f_n\}$ consider $F_i(x) = f_i(x, a, \ldots, a)$, where the first principal subterm of $F_i(x)$ is x and the rest are a. Let $\alpha_i = \alpha(f_i) - 1$, and relabel the f_i if necessary so that $\alpha_n \geq \alpha_{n-1} \geq \cdots \geq \alpha_1$. Define a Knuth–Bendix order on T with precedence $f_n \succ f_{n-1} \succ \cdots \succ f_1 \succ a$ and weights given by:

$$\begin{aligned} w(c) &= 1 \text{ for } c \in \mathcal{F}_C \\ w(x) &= 1 \text{ for } x \in \mathcal{X} \\ w(f_n) &= k \\ w(f_r) &= \mu_r(k + \alpha_n) - \alpha_r \text{ for } r = 1, \ldots, n-1 \end{aligned}$$

where k is chosen large enough that

$$\mu_r k > \alpha_r + 1 \qquad \text{for } r = 1, \ldots, n-1,$$

and so each $w(f_r) > 1$. Now the precedence and the weights satisfy the conditions of the definition, and so they define a Knuth–Bendix order $>_{\mathrm{kbo}}$ on T whose restriction to S can easily be seen to be total. We have

$$\begin{aligned} w(F_r^i(x)) &= i(w(f_r) + \alpha_r) + 1 \\ &= i\mu_r(k + \alpha_n) + 1 \end{aligned}$$

and hence

$$w(F_{r+1}^i(x)) \gtrless w(F_r^j(x)) \quad \text{if} \quad i \gtrless j \frac{\mu_r(k + \alpha_n)}{\mu_{r+1}(k + \alpha_n)}.$$

Thus

$$\tau_r = \frac{\mu_r(k + \alpha_n)}{\mu_{r+1}(k + \alpha_n)}$$

and applying the formula given after Theorem 2.11 we have

$$\tau_{n-1}\cdots\tau_r = \frac{\mu_r(k+\alpha_n)}{\mu_n(k+\alpha_n)} = \mu_r$$

as required. □

Now we investigate one-parameter families for the Knuth–Bendix order in the case $|\mathcal{G}| = 2$. So suppose that $T = T(\mathcal{F})$ and suppose that $S = T_{\mathcal{G},x}$ is a one-parameter family over $\mathcal{G} = \{F, G\}$. We suppose that $>= >_{\mathrm{kbo}}$ is a Knuth–Bendix order on T, defined using the notation of Definition 2.7, and that the restriction of $>$ to S is total, so that, interchanging F and G if necessary, we may assume that $F(x) > G(x)$. It follows from Definition 2.7 that $w(F(x)) = \alpha_F + \rho v_F$ where $\alpha_F \geq 0$ and $v_F = \#(x, F(x))$, the number of occurrences of x in $F(x)$ regarded as an element of T. For $n \geq 1$ we have:

$$\begin{aligned} w(F^n(x)) &= n\alpha_F + \rho \quad \text{if } v_F = 1 \\ &= \alpha_F \frac{(v_F^n - 1)}{(v_F - 1)} + \rho v_F^n \quad \text{if } v_F > 1 \end{aligned}$$

and similarly for G with $w(G(x)) = \alpha_G + \rho v_G$. For clarity of exposition we shall also assume that $\alpha_F, \alpha_G > 0$ and $\alpha_F \neq \alpha_G$. We obtain the following result, which uses the notation of Theorem 2.13 to describe $>_\alpha$.

Theorem 4.8. *Let $T, \mathcal{G}, >, w$ be as above and let $>_\alpha$ be the order induced on S_A by $>$, and $\trianglerighteq_\lambda$ be the λ order with respect to F, G.*
Then one of the following holds:

(1) $v_F = v_G = 1$ and $>_\alpha = \geq_\tau; >$ with $1 > \tau = \frac{\alpha_G}{\alpha_F} \neq 0$. So $>_\alpha$ has:
   - *(a) weight sequence $(1, \tau)$ and*
   - *(b) $F^i(x) \gtrless G^j(x)$ if $i \gtrless j\tau$.*

(2) $v_F > v_G = 1$ and $>_\alpha = \geq_\tau; \trianglerighteq_\lambda; >$ with $\tau = 0$, $\lambda = v_F > 1$ so $>_\alpha$ has:
   - *(a) weight sequence $(1, 0)$ and*
   - *(b) $F(x) > G^i(x)$ for all i and*
   - *(c) $FG^j(x) \gtrless G^k F(x)$ if $jv_F \gtrless k$.*

(3) $v_F \geq v_G > 1$ and $>_\alpha = \geq_\tau; >$ with $1 \geq \tau = \frac{\log v_G}{\log v_F} > 0$ so $>_\alpha$ has:
   - *(a) weight sequence $(1, \tau)$ and*
   - *(b) $F^i(x) \gtrless G^j(x)$ if $i \gtrless j\tau$ and*
   - *(c) $F^i(x) \gtrless G^j(x)$ if $i = j\tau$ and $\frac{\alpha_F}{\alpha_G} \gtrless \frac{v_F - 1}{v_G - 1}$.*

Proof. Note first that $F(x) > G(x)$, so from the definition of the Knuth–Bendix order we have $v_F \geq v_G$, and so we have three cases $v_F = v_G = 1$, $v_F > v_G = 1$ and $v_F \geq v_G > 1$. Furthermore, if $u > v$ then $\#(x, u) > \#(x, v)$, so as $>$ is total on S we have that if

$$\#(x, F^i(x)) \gtrless \#(x, G^j(x)) \text{ then } F^i(x) \gtrless G^j(x),$$

where $\#(x, F^i(x)) = v_F^i$.

(1) If $v_F = v_G = 1$, then $w(F(x)) = \alpha_F + \rho$, $w(G(x)) = \alpha_G + \rho$, and as $F(x) > G(x)$

and $\alpha_F \neq \alpha_G$ we have $\alpha_F > \alpha_G$. Further, $w(F^i(x)) = i\alpha_F + \rho$, $w(G^j(x)) = j\alpha_G + \rho$ so if $\frac{i}{j} \gtrless \frac{\alpha_G}{\alpha_F}$, then $F^i(x) \gtrless G^j(x)$. Thus $>_\alpha = \geq_\tau; >$ with $1 > \tau = \frac{\alpha_G}{\alpha_F} > 0$ and $>_\alpha$ has weight sequence $(1, \tau)$.

(2) If $v_F > v_G = 1$, then for all i, $\#(x, F(x)) = v_F > v_G^i = 1 = \#(x, G^i(x))$, so $F(x) > G^i(x)$. We have $\#(x, FG^j(x)) = v_F = \#(x, G^k F(x))$ and $w(FG^j(x)) = \alpha_F + v_F(j\alpha_G + \rho)$ and $w(G^k F(x)) = k\alpha_G + (\alpha_F + v_F\rho)$, so if $jv_F \gtrless k$, then $w(FG^j(x)) \gtrless w(G^k F(x))$ and thus $FG^j(x) \gtrless G^k F(x)$. Thus $>_\alpha = \geq_\tau; \trianglerighteq_\lambda; >$ with $\tau = 0$ and $\lambda = v_F > 1$, and $>_\alpha$ has weight sequence $(1, 0)$.

(3) If $i \gtrless j\tau$, then $v_F^i \gtrless v_G^j$ and $F^i(x) \gtrless G^j(x)$. If $i = j\tau$ and $\frac{\alpha_F}{\alpha_G} \gtrless \frac{v_F - 1}{v_G - 1}$, then

$$w(F^i(x)) = \alpha_F \frac{v_F^i - 1}{v_F - 1} + \rho v_F^i \gtrless \alpha_G \frac{v_G^j - 1}{v_G - 1} + \rho v_G^j = w(G^j(x)).$$

Thus $>_\alpha = \geq_\tau; >$ with $1 \geq \tau = \frac{\log v_G}{\log v_F} > 0$ and $>$ has weight sequence $(1, \tau)$. □

We have the following corollary.

COROLLARY 4.9. *Let $T, \mathcal{G}, >, w$ be as above, and suppose that at least one of F, G contains more than one occurrence of x. Then the weight sequence for $>_\alpha$ is determined solely by $\#(x, F(x))$ and $\#(x, G(x))$.*

4.3. POLYNOMIAL ORDERS

In this section we investigate polynomial orders in this framework. We suppose that we have a polynomial order on T which induces a total order on a one-parameter family $S = T_{\mathcal{G},x}$, and we compute the values of τ and λ in the case $|\mathcal{G}| = 2$, and the weight sequences in general, from the induced interpretations on the elements of $\mathcal{G}$.

So suppose that $T = T(\mathcal{F} \cup \mathcal{X})$, where $\mathcal{F}_C \neq \emptyset$, and $[.]$ is a polynomial interpretation for $\mathcal{F} \cup \mathcal{X}$ defining an order $>_{\text{poly}}$ on T. Suppose that $S = T_{\mathcal{G},x}$ is a one-parameter family in T, and that $>_{\text{poly}}$ is total on restriction to S and $[.]$ satisfies, for each $F \in \mathcal{G}$, $\deg([F(x)]) > 1$ or $\deg([F(x)]) = 1$ and $[F(x)] = a_F X + b_F$ where $a_F > 1$ or $a_F = 1$ and $b_F > 0$. Since $[s]$ is univariate for each $s \in S$ we may also define $>_{\text{Dpoly}}$ on S.

LEMMA 4.10. *Given $T, [.], \mathcal{G}$ as above then:*

(1) $>_{\text{poly}}|_S = >_{\text{Dpoly}}$ and
(2) $>_{\text{poly}}|_S$ is a simplification order on S, stable under substitution.

PROOF. It follows from the definitions that $>_{\text{poly}}|_S \subseteq >_{\text{Dpoly}}$, as if for some $c \in \mathbb{R}$ we have $p = \sum_{i=1}^m a_i X^i > \sum_{i=1}^n b_i X^i = q$ for all $X > c$, then p eventually dominates q. Thus as $>_{\text{poly}}|_S$ is total we have $>_{\text{poly}}|_S = >_{\text{Dpoly}}$. Part 2 follows from Theorem 3.2 above. □

Then we may use Cropper and Martin (2000) in the case $|\mathcal{G}| = 2$ to extract the values of τ and λ as in Theorem 2.13 to analyse the weight sequences for $>_{\text{poly}}$ in general.

THEOREM 4.11. *Let T, $[.]$, $>_{\text{poly}}$ be as above and suppose that $\mathcal{G} = \{F_1, F_2\}$ where $F_2(x) >_{\text{poly}} F_1(x) >_{\text{poly}} x$ and*

$$[F_1(x)] = b_n X^n + \cdots + b_1 X + b_0$$
$$[F_2(x)] = a_m X^m + \cdots + a_1 X + a_0.$$

Then $>_{\text{poly}}|_S = \geq_\tau; >$ *or* $>_{\text{poly}}|_S = \geq_0; \trianglerighteq_\lambda; >$, *and the weight sequence for* $>_{\text{poly}}|_S$ *is* $(1, \tau)$, *where exactly one of the following holds:*

(1) $m \geq n > 1$ *and* $\tau = \frac{\log n}{\log m}$
(2) $m > n = 1$, $b_1 > 1$, $\tau = 0$ *and* $\lambda = m$
(3) $m > n = 1$, $b_1 = 1$, $b_0 > 0$ *and* $\tau = 0$
(4) $m = n = 1$, $a_1 \geq b_1 > 1$ *and* $\tau = \frac{\log b_1}{\log a_1}$
(5) $m = n = 1$, $a_1 > b_1 = 1$, $b_0 > 0$, $\tau = 0$ *and* $\lambda = a_1$
(6) $m = n = 1$, $a_1 = b_1 = 1$, $a_0 \geq b_0 > 0$, *and* $\tau = \frac{b_0}{a_0}$.

PROOF. From the above discussion $>_{\text{poly}}|_S = >_{\text{Dpoly}}$, and this is just a translation of the main theorem of Cropper and Martin (2000). □

Now we consider the general case, using the values of τ we have just computed.

THEOREM 4.12. *Let* T, $[.]$, $>_{\text{poly}}$, $\mathcal{G}$ *be as above and suppose that* $\mathcal{G} = \{F_1, \ldots, F_n\}$ *has the weight sequence* $1 = \mu_n \geq \mu_{n-1} \geq \cdots \geq \mu_1 \geq 0$, *where*

$$F_n(x) >_{\text{Dpoly}} F_{n-1}(x) >_{\text{Dpoly}} \cdots >_{\text{Dpoly}} F_1(x).$$

Suppose that $[F_i(x)] = k_i X^{m_i} +$ *(terms of lower degree) with* $k_i > 0$, $m_i \geq 1$ *for each* i, *and* $m_n \geq m_{n-1} \geq \cdots \geq m_1$. *Then exactly one of the following holds:*

(1) $m_1 > 1$, $\tau_s = \frac{\log m_s}{\log m_{s+1}}$ *and* $\mu_s = \frac{\log m_s}{\log m_n}$ *for* $s = 1, \ldots, n-1$.
(2) *there is an* r, $n > r \geq 1$, *such that* $m_n \geq \cdots \geq m_{r+1} > m_r = \cdots = m_1 = 1$ *and* $\tau_s = \frac{\log m_s}{\log m_{s+1}}$, $\mu_s = \frac{\log m_s}{\log m_n}$ *for* $n > s \geq r+1$, *and* $\mu_s = 0$ *for* $s = 1, \ldots, r$.
(3) $m_n = \cdots = m_1 = 1$, $k_n \geq \cdots \geq k_1 > 1$ *and* $\tau_s = \frac{\log k_s}{\log k_{s+1}}$ *and* $\mu_s = \frac{\log k_s}{\log k_n} > 0$ *for* $s = 1, \ldots, n-1$.
(4) $m_n = \cdots = m_1 = 1$, *there is an* r, $n > r \geq 1$, *such that* $k_n \geq \cdots k_{r+1} > k_r = \cdots = k_1 = 1$ *and* $\tau_s = \frac{\log k_s}{\log k_{s+1}}$ *and* $\mu_s = \frac{\log k_s}{\log k_n} > 0$ *for* $n > s \geq r+1$, *and* $\mu_s = 0$ *for* $s = 1, \ldots, r$.
(5) $[F_i(x)] = X + p_i$ *for each* i, *with* $p_n \geq \cdots \geq p_1 > 0$ *and* $\tau_s = \frac{p_s}{p_{s+1}}$, $\mu_s = \frac{p_s}{p_n}$ *for* $s = 1, \ldots, n-1$.

PROOF. The five cases come from considering the possible degree sequences of $[F_i(x)]$. In each case we may compute τ_s for $F_{s+1}(x)$, $F_s(x)$ from the previous theorem, and then read off the μ_s from the formula after Theorem 2.11. □

While this result computes the weight sequences given the polynomials interpreting the $F_i(x)$, it does not link these to interpretations for the original function symbols, which will impose further constraints on the $[F_i(x)]$ which we do not consider here in general.

However, we note that if one of the $[F_i(x)]$ is non-linear, then the weight sequence of S is determined by the degrees of the $[F_i(x)]$ alone. If $[f]$ is linear for every function symbol of $\mathcal{F}$ occurring in $\mathcal{G}$, then each $[F_i(x)]$ for $F_i \in \mathcal{G}$ will be linear. Then an argument similar to the one in Theorem 4.7 shows that every possible weight sequence can occur as the weight sequence of a suitably chosen $T(\mathcal{F} \cup \mathcal{X})$ and one-parameter family $S = T_{\mathcal{G},x}$, with $[f]$ linear for each $f \in \mathcal{F}$.

5. Ordering by Counting Patterns

In this section we investigate a general framework for ordering terms by counting patterns. As a simple running example consider terms $T = T(\mathcal{F})$ where $f, g \in \mathcal{F}$ are distinct function symbols with arity at least one. We show that the order on T defined by ordering first on the number of occurrences of g in a term, then on the number of occurrences of an f with a g in the subterm below it, is a partial monotonic order on T. Thus, for $\{a, b, c\}$ constants this order orders $f(a, g(b, c)) > g(f(a, b), f(a, c))$ and $f(g(a, a), g(a, a)) > g(g(a, a), f(a, f(a, a)))$.

In Section 5.1 we describe our general framework for ordering fixed arity terms by counting patterns. We first define the notion of a pattern class P which is the disjoint union of certain subsets P_i of $V(\mathcal{F})$, the set of varyadic terms over $\mathcal{F}$: the definition of P, the set of patterns we are counting, is a certain closure property. In our example $P = \{g, fg\}$, $P_1 = \{g\}$ and $P_2 = \{fg\}$.

We need to count the number of occurrences, that is, distinct embeddings, of patterns in each P_i in an element s of $T(\mathcal{F})$. We record the result $\zeta_i(s)$ as a formal sum over $\mathbb{R}$ of elements of P_i, that is an element of $\mathbb{R}[P_i]$, the real vector space with basis P_i and we define

$$\zeta(s) = \sum_{i \in \mathbb{N}_+} \zeta_i(s) \in (\oplus_i \mathbb{R}[P_i]) = \mathbb{R}[P].$$

Definition 5.9 establishes the ζ_i, ζ and relies on the definition of $\#(w, s)$ given in Subsection 2.1.2.

In our example $\zeta_1(s), \zeta_2(s)$ count the number of embeddings of g, fg respectively in the term s, so that

$$\begin{aligned} \zeta_1(f(a, g(b, c))) &= 1 * (g) \\ \zeta_2(f(a, g(b, c))) &= 1 * (fg) \\ \zeta(f(a, g(b, c))) &= 1 * (g) + 1 * (fg) \\ \zeta(g(f(a, b), f(a, c))) &= 1 * (g) + 0 * (fg). \end{aligned}$$

To define an order on terms we define monotonic orders $>_i$ on each $\mathbb{R}[P_i]$ using the orders on vector spaces we described in Subsection 2.2.4. We then lift their combination $(>_1, >_2, \ldots)$ to an order $>$ on $T(\mathcal{F})$ by

$$s > t \Leftrightarrow \exists j \; \zeta_i(s) = \zeta_i(t) \text{ for } i < j, \; \zeta_j(s) >_j \zeta_j(t).$$

Theorem 5.11 then establishes a key property of ζ which we use in Theorem 5.13 to show that we obtain a simplification order.

Thus in our example if we define $>_1 = >_2 = >_l$, where $>_l$ is the usual linear order on $\mathbb{R} \cong \mathbb{R}[\{g\}] \cong \mathbb{R}[\{fg\}]$, we see that

$$f(a, g(b, c)) > g(f(a, b), f(a, c)).$$

In Section 5.2 we consider three extended examples: we show that even on two unary function symbols our methods give uncountably many new orders, we explain how the Knuth–Bendix order fits into our framework and we sketch the relationship with our earlier zig-zag orders (Martin, 1993).

5.1. THE GENERAL FRAMEWORK

We begin by defining our notion of pattern class and giving some examples. A pattern class consists of the set of patterns we are counting: it needs to satisfy certain extra conditions to make our orders monotonic. Monotonicity is essentially a closure property, and so the intuition behind our definition is that it captures the appropriate closure property for pattern classes for monotonicity to hold.

DEFINITION 5.1. Let $V = V(\mathcal{F})$ be a variable arity term algebra. A *pattern class* consists of a subset P of V together with a decomposition of P as the disjoint union of subsets P_i, $1 \leq i \leq n$ or $1 \leq i \leq \infty$ satisfying:

(1) P_1 is closed under taking principal subterms
(2) if $p \in P_{j+1}$ for some $j \geq 1$ and q is a principal subterm of p, then $q \in Q_j = \cup\{P_i \,|\, 1 \leq i \leq j\}$.

Thus, in our running example we would take $P = \{g, fg\}$, $P_1 = \{g\}$ and $P_2 = \{fg\}$: then P_1, P_2 and $Q_2 = P$ have the required properties.

EXAMPLE 5.2. If P is any subset of V closed under taking principal subterms, then P can be regarded as a pattern class if we take P_i as the elements of P of height i, so Q_i is the elements of height at most i.

EXAMPLE 5.3. In particular if P and Q are closed under taking principal subterms, then so are $P \cap Q$ and $P \cup Q$, which may thus also be regarded as pattern classes.

EXAMPLE 5.4. Thus, in particular if $\mathcal{G} \subseteq \mathcal{F}$, then $P = V(\mathcal{G})$ or $P = T(\mathcal{G})$ may be regarded as pattern classes by taking P_i to be the elements of P of height i, so that Q_i is the elements of height at most i.

EXAMPLE 5.5. Let P be a subset of V closed under homeomorphic embedding for variable arity terms, so in particular P is closed under taking principal subterms, and suppose that P is the disjoint union of subsets P_i for $1 \leq i \leq n$ or $1 \leq i \leq \infty$ where each P_i is an antichain (that is to say it contains no comparable pairs of elements) and each $Q_j = \dot\cup_{i=1}^{j} P_i$ is closed under varyadic embedding. Then P is a pattern class as each Q_j is closed under taking principal subterms and if $p \in P_{j+1}$ and q is a principal subterm of p, then q embeds in p, so by the antichain property $q \in Q_{j+1} \setminus P_{j+1} = Q_j$.

EXAMPLE 5.6. If $P \subseteq V$ is closed under homeomorphic embedding for variable arity terms $\succeq_{\text{emb}}$, then we may regard P as a pattern class if we take

$$P_1 = \min_{\succeq_{\text{emb}}}(P)$$
$$P_{i+1} = \min_{\succeq_{\text{emb}}}(P \setminus P_i) \qquad \text{for } i \geq 1$$

EXAMPLE 5.7. As an instance of Example 5.5 take X to be any subset of $V = V(\mathcal{F})$ and P the set of all terms in $V(\mathcal{F})$ or $T(\mathcal{F})$ with at most one occurrence of any element of X as an embedded subterm. Thus, for example if $X = \{a, b\}$, then each term in P, contains no as or bs at all, or one a or one b.

EXAMPLE 5.8. As a further instance of Example 5.5 take X to be any finite subset of $V(\mathcal{F})$ and let

$$P = Forb(X) = \{s \in V(\mathcal{F}) | x \in X \Rightarrow \neg(x \,{}_{emb}\!\prec s)\}$$

(and likewise for $T(\mathcal{F})$). P is closed under embedding.

If H is a set, then the set of formal sums $\mathbb{R}[H]$ of H over $\mathbb{R}$ is defined as the vector space over $\mathbb{R}$ with basis H. Thus we may define the usual vector space operations: for example the direct sum $\mathbb{R}[P] \oplus \mathbb{R}[Q]$ may be identified with the formal sum $\mathbb{R}[P \cup Q]$.

Now let $\mathbb{R}[P] = \oplus_{i=1}\mathbb{R}[P_i]$ be the set of formal sums over $\mathbb{R}$ of elements of P. We define the map

$$\zeta : T(\mathcal{F}) \to \mathbb{R}[P]$$

as the direct sum of the component maps

$$\zeta_i : T(\mathcal{F}) \to \mathbb{R}[P_i]$$

defined below. Informally, $\zeta_i(s)$ "counts how many times each element of P_i embeds in s".

DEFINITION 5.9. Let $P = \dot\cup P_i$ be a pattern class and $s \in T(\mathcal{F})$. We define ζ_i, for $i \geq 1$ as

$$\zeta_i(s) = \left(\sum_{u \in P_i} \#(u, s)u \right) \in \mathbb{R}[P_i].$$

We say that w *occurs in* s if $w \in P$ and $\#(w, s) \neq 0$

Our next task is to compute the ζ_i, by counting the occurrences of elements w of P in a term s. In the following theorem, $\zeta_i(s)$ will be split into three components, which correspond to the three structures of terms which w can take in the definition of $\#(w, s)$ in Subsection 2.1.2. We therefore introduce the following notation.

NOTATION 5.10. Let $P = \dot\cup P_i$ be a pattern class, and let $s = f(s_1, \ldots, s_n) \in T(\mathcal{F})$. Then let

$$\begin{aligned}
\mathcal{F}_i &= \{w \in P_i | w \in \mathcal{F}\} \\
\alpha_{i,s} &= \{w \in P_i | (\mathrm{hd}(w) \neq f) \vee ((\mathrm{hd}(w) = f) \wedge (\mathrm{hdar}(w) > n))\} \\
\beta_{i,s} &= \{w \in P_i |\, \mathrm{hd}(w) = f \wedge n \geq \mathrm{hdar}(w)\}.
\end{aligned}$$

The following result shows that ζ has the key property which will ensure that our orders are monotonic.

THEOREM 5.11. *Let $P = \dot\cup P_i$ be a pattern class, and let*

$$\zeta = \oplus_i \zeta_i : T(\mathcal{F}) \to \mathbb{R}[P_i]$$

be as above. Let $s, t, u = f(s, s_2, \ldots, s_n)$, and $v = f(t, s_2, \ldots, s_n)$ be elements of $T(\mathcal{F})$, and let $k > 1$. Suppose $\forall h < k$, $\zeta_h(s) = \zeta_h(t)$. Then $\forall r \leq k$

$$\zeta_r(s) - \zeta_r(t) = \zeta_r(u) - \zeta_r(v).$$

PROOF. First note that we can decompose $\zeta_d(u)$ for any $d \geq 1$ in the following way:

$$\begin{aligned}\zeta_d(u) &= \sum_{w \in P_d} \#(w,u)w \\ &= \sum_{w \in \mathcal{F}_d} Aw + \sum_{w \in \alpha_{d,u}} Bw + \sum_{w \in \beta_{d,u}} Cw\end{aligned}$$

where

$$A = \#(w,f) + \#(w,s) + \sum_{i=2}^{n} \#(w,s_i)$$

$$B = \#(w,s) + \sum_{i=2}^{n} \#(w,s_i), \text{ and for } w = f(w_1,\ldots,w_m) \in \beta_{d,u} (n \geq m),$$

$$C = \#(w,s) + \sum_{i=2}^{n} \#(w,s_i) + \#(w_1,s) \sum_{(j_k)_{k=2}^m \sqsubset [2,n]} \prod_{i=2}^{m} \#(w_i,s_{j_i})$$

$$+ \sum_{(j_k)_{k=1}^m \sqsubset [2,n]} \prod_{i=1}^{m} \#(w_i,s_{j_i}).$$

Also note that $\mathrm{hd}(u) = f = \mathrm{hd}(v)$ and $\mathrm{hdar}(u) = n = \mathrm{hdar}(v)$. It follows that $\alpha_{d,u} = \alpha_{d,v}$ and $\beta_{d,u} = \beta_{d,v}$ for each $d \in \mathbb{N}_+$.

Now consider the case $r = 1$. Then $r = 1 < k$ so $\zeta_1(s) = \zeta_1(t)$ and so

$$\sum_{w \in P_1} \#(w,s)w = \sum_{w \in P_1} \#(w,t)w.$$

In particular, as P_1 is a basis for $\mathbb{R}[P_1]$, for all $w \in P_1$ we have $\#(w,s) = \#(w,t)$. Hence

$$\zeta_1(u) - \zeta_1(v) = \sum_{w \in \beta_{1,u}} (\#(w_1,s) - \#(w_1,t)) \sum_{(j_k)_{k=2}^m \sqsubset [2,n]} \prod_{i=2}^{m} \#(w_i,s_{j_i})w.$$

Now we have $w_1 \in P_1$ as P_1 is closed under taking principal subterms. So $\#(w_1,s) = \#(w_1,t)$ and thus $0 = \zeta_1(u) - \zeta_1(v) = \zeta_1(s) - \zeta_1(t)$, as required.

Now consider the case $r > 1$. So, $\forall h < k$, $\zeta_h(s) = \zeta_h(t)$; i.e.

$$\sum_{w \in P_h} \#(w,s)w = \sum_{w \in P_h} \#(w,t)w$$

and hence $\#(w,s) = \#(w,t)$ for each $w \in P_h$. Now consider $1 < r \leq k$. It follows from the decomposition of the formal sum given above that:

$$\zeta_r(u) - \zeta_r(v) = \zeta_r(s) - \zeta_r(t) + \sum_{w \in \beta_{r,u}} (\#(w_1,s) - \#(w_1,t)) \sum_{(j_k)_{k=2}^m \sqsubset [2,n]} \prod_{i=2}^{m} \#(w_i,s_{j_i})w.$$

Now consider w_1. Since $Q_r = \dot\bigcup_{j=1}^{r} P_j$ is closed under taking principal subterms we have $w_1 \in Q_r$. As $w \in P_r$, the second condition in the definition of pattern class states that w_1 is in $Q_{r-1} = \dot\bigcup_{j=1}^{r-1} P_j$. Hence $w_1 \in P_h$ for some $h < r \leq k$; i.e. for $h \leq k-1$. Thus $\#(w_1,s) = \#(w_1,t)$ and so $\zeta_r(u) - \zeta_r(v) = \zeta_r(s) - \zeta_r(t)$ as required. □

Now we can construct the orders we seek.

Definition 5.12. Given $\mathcal{F}$, P and ζ as above, a *monotonic family* for P consists of a sequence $(>_i)_{i\geq 1}$, where each $>_i$ is a monotonic order on the real vector space $\mathbb{R}[P_i]$ (see Subsection 2.2.4), and given $s, t, u = f(\ldots, s_{i-1}, s, s_{i+1}, \ldots), v = f(\ldots, s_{i-1}, t, s_{i+1}, \ldots) \in T(\mathcal{F})$ we have that $\zeta_1(s) >_1 \zeta_1(t)$ implies $\zeta_1(u) >_1 \zeta_1(v)$.

A monotonic family $(>_i)$ for P defines a relation $>$ on $T(\mathcal{F})$ by

$$\begin{aligned} s > t &\Leftrightarrow \zeta(s)(>_1, >_2, \ldots)\zeta(t) \\ &\Leftrightarrow \exists j \geq 1 \text{ such that } \zeta_i(s) = \zeta_i(t) \text{ for each } i < j \text{ and } \zeta_j(s) >_j \zeta_j(t). \end{aligned}$$

We say $P, (>_i)$ distinguish s, t in $T(\mathcal{F})$ if either $s > t$ or $t > s$.

In our running example, $(>_1, >_2)$ is a monotonic family as each $>_i$ is monotonic on $\mathbb{R}[P_i]$ and ζ_1 satisfies the additional condition as it merely counts the number of occurrences of the letter g in a term.

Finally the next theorem shows that monotonic families over P define monotonic orders on terms.

Theorem 5.13. *Let $P = \dot\cup P_i$ be a pattern class in $V(\mathcal{F})$, ζ be defined as above and $>= (>_i)$ a monotonic family for P. Then $>$ is a monotonic strict partial order on $T(\mathcal{F})$.*

Proof. We need to show that $>$ is monotonic, that is, that if $s > t$ then $f(s_1, \ldots, s_n) > f(t_1, \ldots, t_n)$ whenever there is a p $(1 \leq p \leq n)$ with $s_p = s$, $t_p = t$ and $s_i = t_i$ for $i \neq p$. Without loss of generality it suffices to consider the case $p = 1$. Suppose that $s, t, u = f(s, s_2, \ldots, s_n), v = f(t, s_2, \ldots, s_n) \in T(\mathcal{F})$ and $s > t$. Then there is a k such that $\zeta_k(s) >_k \zeta_k(t)$ and $\zeta_h(s) = \zeta_h(t)$ for $h < k$.

If $k = 1$, then $\zeta_1(s) >_1 \zeta_1(t)$ implies that $\zeta_1(u) >_1 \zeta_1(v)$ by hypothesis. Hence $u > v$ by definition of $>$. If $k > 1$, then by Theorem 5.11, for all $r \leq k$ we have

$$\zeta_r(u) - \zeta_r(v) = \zeta_r(s) - \zeta_r(t).$$

Thus if $r < k$ we have

$$\zeta_r(u) - \zeta_r(v) = 0$$

and for $r = k$ we have

$$\zeta_k(u) - \zeta_k(s) = \zeta_k(v) - \zeta_k(t)$$

and hence as $>_k$ is monotonic

$$(\zeta_k(u) - \zeta_k(s)) + \zeta_k(s) >_k (\zeta_k(v) - \zeta_k(t)) + \zeta_k(t)$$

and so $\zeta_k(u) >_k \zeta_k(v)$. Thus $u > v$ as required. □

The framework we have established allows very general classes of pattern classes P and monotonic ordering families $(>_i)$: we are also interested in the case when they give rise to simplification orders. We note that:

(1) The order we construct may fail to have the subterm property, or be total, either because P is "too small" so $\zeta(s) = \zeta(t)$ for $s \neq t$, or because P is "large enough" but the orders $(>_i)$ do not order enough terms, so $\zeta_i(s)$ and $\zeta_i(t)$ are distinct and incomparable for some s, t with $\zeta_r(s) = \zeta_r(t)$ for $r < i$.
(2) Suppose that t is a subterm of a term s: then $\zeta(s) - \zeta(t)$ will always be of the form $\sum_{p \in P} r\ p$ where each $r \geq 0$. If we require that P be large enough so that for each

such pair s, t, then at least one $r \neq 0$, and further that the $(>_i)$ be strong enough that for each $p \in P_i,\ p >_i \underline{0} \in \mathbb{R}[P_i]$, then $>$ will have the subterm property and will thus be a simplification order and well-founded. To ensure the former condition we would need to formalize some notion of a set of varyadic terms forming a basis for terms. To ensure the latter it is sufficient (but not necessary) for each $>_i$ to be total, so defined by orthogonal matrices as indicated in Subsection 2.2.4.

(3) To attempt to construct an order of this kind which proves a particular rewrite system terminating we need to identify the occurrences of varyadic terms embedding into each side of each of the rules, and then to construct a suitable pattern class and order. Example 5.6 shows how we may take for P the closure under varyadic embedding of all subterms of all of the rules. We then need to try to construct orders on each $\mathbb{R}[P_i]$ which make each left-hand side greater than each right-hand side: by the theory of orders on vector spaces indicated in Subsection 2.2.4 this reduces to solving real linear inequalities. We have yet to carry out a practical investigation of the technique: it would have many similarities with the technique for using the simplex method to identify weights for the Knuth–Bendix ordering (Dick *et al.*, 1990).

In our running example the order we have given is not strong enough to distinguish between terms and subterms, for example $f(f(a, a), a)$ and $f(a, a)$.

We return to the example of Section 1, and use the pattern class $P = \{g, g(g)\}$, $P_1 = \{g\}$, $P_2 = \{g(g)\}$, and the orders $>_1 = >_2 = >_l$. This defines a partial monotonic order on $T(\mathcal{F})$ for any g in $T(\mathcal{F})$. We have for example that

$$g(a, g(a, g(a, a))) > g(g(a, a), g(a, a))$$

as

$$\zeta(g(a, g(a, g(a, a)))) = 3 * (g) + 3 * (g(g))$$
$$\zeta(g(g(a, a), g(a, a))) = 3 * (g) + 2 * (g(g)).$$

If $\mathcal{F} = \{g\} \cup \mathcal{F}_C$, then this order always differentiates between terms and subterms, so it is a partial simplification order.

If we take $P = \{g, g(g, g)\}$ and $>_1 = >_2 = >_l$, then we obtain

$$g(g(a, a), g(a, a)) > g(a, g(a, g(a, a)))$$

as

$$\zeta(g(a, g(a, g(a, a)))) = 3 * (g) + 0 * (g(g, g))$$
$$\zeta(g(g(a, a), g(a, a))) = 3 * (g) + 1 * (g(g, g)).$$

Applying the method indicated above in (3) we would attempt to build an order by considering $P = \{g, g(g), g(g, g), \ldots\}$ (suppressing some elements of P for clarity) with $P_1 = \{g\}, P_2 = \{g(g)\}, P_3 = \{g(g, g)\}, \ldots$, and on each P_i would have a choice of orders $>_i$ to construct the order we need: in this case as each of P_1, P_2, P_3 has dimension 1 the theory of Subsection 2.2.4 shows that we have a choice of three orders for each $>_i$, and suitable choices give us the two orders above and thus a choice for ordering $g(a, g(a, g(a, a)))$ and $g(g(a, a), g(a, a))$.

5.2. Examples

In this section we present some simple examples to indicate the power of these methods. We then show how these orders subsume the weight part of the Knuth–Bendix order and the zig-zag order of Martin (1993).

Example 5.14. We consider terms $T = T(\{f, g\} \cup \{a\})$ in two unary operators f, g and a constant a. We let

$$P_1 = \{f(a), g(a), a\}$$
$$P_2 = \{fg(a), gf(a)\}$$
$$\text{and } P_3 = \{fgf(a), gfg(a)\}$$

and it is easy to see that $P = P_1 \cup P_2 \cup P_3$ is a pattern class. Let $>_1$, $>_2$ and $>_3$ be monotonic orders on $\mathbb{R}[P_1]$, $\mathbb{R}[P_2]$ and $\mathbb{R}[P_3]$ respectively and $>$ the order they induce on T. We show that for $>_1$ and $>_2$ fixed, there are continuum many choices for $>_3$, giving continuum many distinct orders $>$. To see this consider

$$s = f^N g^{2M} f^N(a) \quad \text{and} \quad t = g^M f^{2N} g^M(a).$$

It is clear that

$$\zeta_1(s) = \zeta_1(t) = 2N\ f(a) + 2M\ g(a) \text{ and}$$
$$\zeta_2(s) = \zeta_2(t) = 2NM\ (fg(a) + gf(a)).$$

We have

$$\zeta_3(s) = 2MN^2\ fgf(a) \text{ and}$$
$$\zeta_3(t) = 2M^2N\ gfg(a).$$

Now for any positive real λ define $>_{3,\lambda}$ by

$$\alpha\ fgf(a) + \beta\ gfg(a) >_{3,\lambda} \alpha'\ fgf(a) + \beta'\ gfg(a) \Longleftrightarrow \alpha + \beta\ \lambda > \alpha' + \beta'\ \lambda.$$

Then

$$\zeta_3(s) >_{3,\lambda} \zeta_3(t)$$

if and only if

$$2MN^2 > 2M^2N\ \lambda$$

that is

$$N/M > \lambda$$

and

$$\zeta_3(s)\ {}_{3,\lambda}{<}\ \zeta_3(t)$$

if and only if

$$2MN^2 < 2M^2N\ \lambda \qquad \text{that is } N/M < \lambda.$$

Define $>_\lambda$ to be the lifting of $(>_1, >_2, >_{3,\lambda})$ to T extended to a total order by comparing lexicographically from the left with $f \succ g$. We have the following.

Theorem 5.15. *Let $>_1$, $>_2$, $>_{3,\lambda}$ and $>_\lambda$ be as above. Then:*

(1) For any monotonic $>_1$, $>_2$, total $>_1$ and real positive λ, $>_\lambda$ is a total simplification order.
(2) For fixed $>_1$, $>_2$, $\lambda_1 \neq \lambda_2$, then $>_{\lambda_1} \neq >_{\lambda_2}$.

PROOF. (1) Is just an application of the results from Section 5.1.
(2) Suppose without loss of generality that $\lambda_1 > \lambda_2 > 0$. Then there is a rational number N/M such that $\lambda_1 > N/M > \lambda_2$. But then by the argument above we have

$$f^N g^{2M} f^N(a) >_{\lambda_2} g^M f^{2N} f^M(a) \quad \text{and} \qquad f^N g^{2M} f^N(a) \;{}_{\lambda_1}\!< g^M f^{2N} f^M(a)$$

and so $>_{\lambda_1} \neq >_{\lambda_2}$. $\square$

EXAMPLE 5.16. We show how the weight pre-order of the Knuth–Bendix order can be obtained by a pattern order.

Let $\mathcal{F}$ be a set of function symbols. Then it is easy to check that $P = P_1 = \mathcal{F}$ is a pattern class. Therefore, for any $t \in T(\mathcal{F})$, $\zeta(t) = \sum_{f \in \mathcal{F}} \#(f, t) f$; i.e. ζ counts the multiplicities of the function symbols in t. It follows that $\mathbb{R}[P] = \mathbb{R}[P_1]$ is an $\mathbb{R}$ vector space with basis elements of $\mathcal{F}$ and dimension $|\mathcal{F}|$, the cardinality of $\mathcal{F}$. By the comments in Subsection 2.2.4, we can assign a weight vector $w \in \mathbb{R}^{|\mathcal{F}|}$ and define, for $u, v \in T(\mathcal{F})$, $u > v$ if and only if $\zeta(u).w > \zeta(v).w$, where . is the usual dot product of vectors. We therefore obtain the pre-order by weight of the Knuth–Bendix order.

We may extend this order in continuum many distinct ways by counting additional patterns as in Theorem 5.15.

EXAMPLE 5.17. We briefly discuss the "zig-zag" orders of Martin (1993). Zig-zags are annotated sequences of function symbols. Formally:

DEFINITION 5.18. A *zig-zag* $Z = (z, \eta)$ of length n on a set of function symbols $\mathcal{F}$ consists of a non-empty sequence $z = [f_1, \ldots, f_n]$ of elements of $\mathcal{F}$ and a total function $\eta : [1, \ldots, n-1] \to \mathbb{N}$ such that $1 \leq \eta(i) \leq \alpha(f_i)$ for each $1 \leq i \leq n-1$. For convenience Z is represented as a labelled directed graph

$$f_1 \stackrel{\eta(1)}{\to} f_2 \stackrel{\eta(2)}{\to} \cdots f_{n-1} \stackrel{\eta(n-1)}{\to} f_n.$$

We denote by $Z_{\mathcal{F}}(n)$ or $Z(n)$ the set of all zig-zags of length n on $\mathcal{F}$.

Zig-zags represent paths in a term. For example, $f \stackrel{2}{\to} g \stackrel{1}{\to} a$ and $f \stackrel{1}{\to} g$ are zig-zags on $\mathcal{F} = \{f, g, a\}$ where f and g have arity 2 and a has arity 0, and similarly we can represent $f(g(a, a), g(a, a))$ contains one occurrence each of $f \stackrel{2}{\to} g \stackrel{1}{\to} a$ and $f \stackrel{1}{\to} g$. In Martin (1993) simplification orders are constructed on $T(\mathcal{F})$ by mapping a term s to the formal sum of the zig-zags it contains, by $\zeta : T(\mathcal{F}) \to \oplus\mathbb{R}[Z_{\mathcal{F}}(n)]$, and then using ζ to lift appropriate monotonic families $>= (>_i)$ of orders $>_i$ on each $\mathbb{R}[Z_{\mathcal{F}}(i)]$ to orders on $T(\mathcal{F})$.

Our objective is to define a pattern class on $V(F)$ which behaves like the set of zig-zags $Z_{\mathcal{F}}(n)$. However, if $f, g \in \mathcal{F}$ have arity greater than 1 there is no satisfactory way for elements in $V(\mathcal{F})$ to distinguish between $f \stackrel{1}{\to} g$ and $f \stackrel{2}{\to} g$ in $Z_{\mathcal{F}}(n)$. Instead, we introduce a constant symbol $\square$ and define a pattern class P on $T(\mathcal{F} \cup \{\square\})$ which behaves as required. Here $\square$ will represent a "redundant" subterm. For example, if f and g have arity 2 we can represent $f \stackrel{1}{\to} g$ by $f(g(\square, \square), \square)$; i.e. g occurs in the first subterm of f, the second subterm of f is redundant and the subterms of g are redundant, and similarly we can represent $f \xrightarrow{2} g$ by $f(\Box, g(\Box, \Box))$. We now define $P = \dot{\cup} L^n$ in $T = T(\mathcal{F} \cup \{\Box\})$.

DEFINITION 5.19. Let $T = T(\mathcal{F} \cup \{\Box\})$ be a term algebra. Then let $T_\Box$, the *zig-zag* terms of T, be defined by:

(1) $\mathcal{F}_C \cup \{\Box\} \subseteq T_\Box$ and
(2) $f(s_1, \ldots, s_n) \in T_\Box$ if and only if there is an s_j, such that $s_j \in T_\Box$ and $s_i = \Box$ for $i \neq j$.

Let $T^i_\Box$ denote the set of terms of $T_\Box$ of height i. Let $U^i_\Box$ denote the set of terms of $T_\Box$ of height i in which $\Box$ is the only constant symbol, and let L^n denote the set $(T^n_\Box \backslash U^n_\Box) \cup U^{n+1}_\Box$.

It is easy to check that $P = \dot{\cup} L^n$ is a pattern class. We can now define, for each n, a mapping γ_n from L^n to $Z(n)$.

DEFINITION 5.20. Let $\mathcal{F}$ be a set of function symbols where each $f_i \in \mathcal{F}$ has arity α_i, $\Box$ be a constant, $T = T(\mathcal{F} \cup \{\Box\})$ be a term algebra with $T_\Box$ being a set of zig-zag terms of T and Z be a set of zig-zags on $\mathcal{F}$. For each $n \geq 1$ we define the function

$$\gamma_n : L^n \rightarrow Z(n)$$

as follows. For $q \in L^1$ we have $\gamma_1(q) = f$ where f is the head of q. For $q = f(\Box, \ldots, q_i, \ldots, \Box)$ in L^n where $q_i \neq \Box$ is the ith argument of f we have $\gamma_n(q) = f \xrightarrow{i} \gamma_{n-1}(q_i)$.

We then have the following theorem.

THEOREM 5.21. *Let $L^n, Z(n), \gamma_n$ be defined as above. Then γ_n is a bijection for each $n \geq 1$.*

PROOF. Follows by induction on n. $\Box$

To complete our construction of the zig-zag order via patterns we need to define a slight variant of the ζ of Definition 5.9, since terms in $T(\mathcal{F})$ do not contain any occurrence of $\Box$. We do this by extending the base case of Definition 2.4 so that $\#(\Box, s) = 1$ for any term s, and hence $\#(u, s) \geq 1$ for each u in P. We can then define ζ in terms of this extended definition of $\#$, and obtain the following

THEOREM 5.22. *Let $P = \dot{\cup} L^i$, and*

$$\zeta = \oplus_i \zeta_i : T(\mathcal{F}) \rightarrow \mathbb{R}[L^i]$$

be as above. Let $s, t, u = f(s, s_2, \ldots, s_n)$, and $v = f(t, s_2, \ldots, s_n)$ be elements of $T(\mathcal{F})$, and let $k > 1$. Suppose $\forall h < k$, $\zeta_h(s) = \zeta_h(t)$. Then $\forall r \leq k$

$$\zeta_r(s) - \zeta_r(t) = \zeta_r(u) - \zeta_r(v).$$

PROOF. Similar to Theorem 5.11. $\Box$

It follows that ζ allows us to lift suitable orders on the L^i to orders on $T(\mathcal{F})$, and we can use the bijection of Theorem 5.21 to show that these simulate the zig-zag orders of Martin (1993).

6. Patterns for the Recursive Path Order

In this section we consider how our ideas of enumeration in the previous section may be modified to deal with the recursive path order. We note first that we cannot extend our previous results directly.

Example 6.1. We construct a term s with the property that for any p which strictly embeds in s, there is a term t with $s >_{\text{rpo}} t$ but p occurring more often in t than in s. This means that $>_{\text{rpo}}$ cannot be induced by a pattern order.

Let $\{f, a\} \subseteq \mathcal{F}$, where f has arity 2 and a is a constant, and consider the terms

$$s = f(a, f(a, f(f(a, a), f(a, a))))$$
$$t = f(f(k, k), f(k, k)) \text{ where } k = f(f(a, a), a).$$

If p embeds in s, then either p embeds in t, and occurs more times in t than in s, or p embeds in t' where t' is got from t by replacing k by $f(a, f(a, a)$ and again p occurs more times in t' than s. The proof is by exhaustion. We have $s >_{\text{rpo}} t$ and $s >_{\text{rpo}} t'$ as required.

However, under certain mild restrictions on the precedence we can characterize the recursive path order by means of the skeleton of a term s. Suppose that $\mathcal{F}$ is a finite set of function symbols and that $\mathcal{F}_C \subset \mathcal{F}$ is non-empty. Suppose that $\succ$ is a partial order on F, total on F_C, and that $\mathcal{F} \setminus \mathcal{F}_C$ has a greatest element f. It follows that any term s contains a maximal embedded subterm on $\mathcal{G} = \{f\} \cup \mathcal{F}_C$: we call this subterm the skeleton of s, denoted skel(s). In Theorem 6.5 we show that the recursive path order defined by $\succ$ on $T(\mathcal{F})$ induces the height pre-order on the skeletons, so that if $\text{height}(\text{skel}(s)) > \text{height}(\text{skel}(t))$, then $s >_{\text{rpo}} t$. Thus, for example, if $s = f(f(a, a), a)$, then $s = \text{skel}(s)$ and $\text{height}(\text{skel}(s)) = 2$, so s is greater than any term t containing at most one occurrence of f, since $\text{height}(\text{skel}(t)) < 2$.

It is tempting to conjecture a stronger result, that if $\text{skel}(s) >_{\text{rpo}} \text{skel}(t)$, then $s >_{\text{rpo}} t$. This is true for unary function symbols, but we may exhibit s, t with $s >_{\text{rpo}} t = \text{skel}(t) >_{\text{rpo}} = \text{skel}(s)$: take $s = f(g(f(a, a), a))$, so $\text{skel}(s) = f(f(a, a), a)$, and $t = \text{skel}(t) = f(f(a, a), f(a, a))$.

In particular, if $\mathcal{F} = \{f\} \cup \mathcal{F}_C$, then skeleton of s is just s itself. In Theorem 6.10 we show that in this case the recursive path order is just an extension of an order by height by a further order, which compares two terms of equal height p by comparing the largest subterms of height q at depth $p - q$ for some value of q.

If further $\mathcal{F}_C$ contains just one element, then Corollary 6.11 shows how this result may be simplified by comparing certain maximal subtrees of fixed height. Theorem 6.12 identifies a related phenomenon in the case of monadic terms.

6.1. skeletons

In this section we define the f-skeleton and prove Theorem 6.5.

Definition 6.2. Let $\mathcal{F}$ be a finite set of function symbols, and suppose that $\mathcal{F}_C \subset \mathcal{F}$ is non-empty. Suppose that $\succ$ is a partial order on $\mathcal{F}$, total on $\mathcal{F}_C$, and that $\mathcal{F} \setminus \mathcal{F}_C$ has a greatest element f so that $f \succ g$ for all $f \neq g$ in $\mathcal{F}$. Let $\mathcal{G} = \{f\} \cup \mathcal{F}_C$. Let $s \in T(\mathcal{F})$ and let

$$\hat{s} = \{t \in T(\mathcal{G}) | t \;{}_{emb}\!\prec s\}.$$

Note that $\hat{s}$ will be non-empty as it will at contain at least one constant. It follows from our hypotheses that $\succ$ is total on $\mathcal{G}$, so $\geq_{\text{rpo}}$ is total on $\hat{s}$, and so there is an element q of $\hat{s}$ such that $q \geq_{\text{rpo}} p$ for all $p \in \hat{s}$, and q is unique up to $\geq_{\text{rpo}}$ equivalence. We define $\text{skel}(s)$ to be some embedded subterm of s equivalent to q.

Example 6.3. Let $\mathcal{F} = \{f = f_1, \ldots, f_n, a\}$ where each f_i $(1 \leq i \leq n)$ is unary, f is the greatest element of $\mathcal{F}$ and a is a constant. For any $s \in T(\mathcal{F})$, $\text{skel}(s) = f^n(a)$ where $n = \#(f, s)$.

Example 6.4. Let $\mathcal{F} = \{f, g, a\}$ where a is a constant and f and g have arity 2, and let $f \succ g \succ a$. Let

$$s = f(g(f(f(a,a),a), f(a, f(a,a))), g(a,a))$$

and

$$t = g(f(a,a), f(f(a,a), g(a,a)))$$

be elements of $T(\mathcal{F})$. Then

$$\text{skel}(s) = f(f(f(a,a),a),a) \approx_{\text{rpo}} f(f(a, f(a,a)),a)$$

as $f(f(a,a),a) \approx_{\text{rpo}} f(a, f(a,a))$, and $\text{skel}(t) = f(f(a,a),a)$.

We can now prove our main result.

Theorem 6.5. *Let $f, \mathcal{F}$ and $\succ$ be as above. Then, for $s, t \in T(\mathcal{F})$,*

(1) if $\text{height}(\text{skel}(s)) > \text{height}(\text{skel}(t))$, *then* $s >_{\text{rpo}} t$
(2) if $s \geq_{\text{rpo}} t$, *then* $\text{height}(\text{skel}(s)) \geq \text{height}(\text{skel}(t))$.

Proof. (1) We proceed by induction on $\text{height}(t)$.

Case $\text{height}(t) = 1$**.** Then $t \in \mathcal{F}_C$, i.e. t is a constant. Then $\text{height}(t) = \text{height}(\text{skel}(t)) = 1 < \text{height}(\text{skel}(s))$. So s contains an f, and by the definition of the precedence $\succ$ on $\mathcal{F}$, $f \succ t$, and hence $s >_{\text{rpo}} t$.
Case $\text{height}(t) > 1$. Let $s = h(s_1, \ldots, s_m)$ and $t = g(t_1, \ldots, t_n)$. Without loss of generality, we may assume that $h = f$ by replacing s with the smallest subterm of s containing $\text{skel}(s)$: $\text{skel}(s)$ must contain at least one occurrence of f as $\text{height}(\text{skel}(s)) > \text{height}(\text{skel}(t))$. We now have two cases to consider:

Case $g \neq f$**.** It follows that $\text{skel}(t) = \text{skel}(t_i)$ for some i, and so $\text{height}(\text{skel}(s)) > \text{height}(\text{skel}(t)) = \text{height}(\text{skel}(t_i)) \geq \text{height}(\text{skel}(t_j))$ for $j \neq i$. So for $1 \leq j \leq n$, $s >_{\text{rpo}} t_j$ by induction and $f \succ g$ by hypothesis. Hence $s >_{\text{rpo}} t$.
Case $g = f$**.** It follows that

$$\begin{aligned} \text{height}(\text{skel}(s)) &= 1 + \max\{\text{height}(\text{skel}(s_i))\} \\ &> \text{height}(\text{skel}(t)) \\ &= 1 + \max\{\text{height}(\text{skel}(t_j))\}. \end{aligned}$$

Therefore, there is an i such that for all j height(skel(s_i)) $>$ height(skel(t_j)) and hence $s_i >_{\rm rpo} t_j$ for $1 \leq j \leq n$ by induction. Thus $s >_{\rm rpo} t$.

(2) Suppose that $s >_{\rm rpo} t$ and height(skel(s)) $<$ height(skel(t)). Then by part 1, $t >_{\rm rpo} s$ which is a contradiction, and part 2 follows immediately. □

Example 6.3 gives an immediate corollary in the case of monadic terms, confirming what we already know from Martin and Scott (1997). We have the following.

Corollary 6.6. *Let $\mathcal{F} = \{f_1, \ldots, f_n, a\}$ where each f_i $(1 \leq i \leq n)$ is unary, a is a constant, $\succ$ is total and $f_1 \succ f_j$ for $2 \leq j \leq n$. Then:*

(1) if $\#(f_1, s) > \#(f_1, t)$, then $s >_{\rm rpo} t$
(2) if $s >_{\rm rpo} t$, then $\#(f_1, s) \geq \#(f_1, t)$.

Further consideration of this example shows that an obvious generalization of Theorem 6.5 is false. Suppose $f_1 \succ f_2 \succ f_3$ and we define the $\{f_1, f_2\}$-skeleton denoted by $\mathrm{skel}_{\{f_1,f_2\}}(s)$ in the obvious way. Then it is not true that

$$\mathrm{height}(\mathrm{skel}_{\{f_1,f_2\}}(s)) > \mathrm{height}(\mathrm{skel}_{\{f_1,f_2\}}(s)) \Rightarrow s >_{\rm rpo} t.$$

For consider $s = f_1 f_2 f_3(a), t = f_2^2 f_1 f_3 f_2(a)$. Then $s >_{\rm rpo} t$ but $\mathrm{skel}_{\{f_1,f_2\}}(s) = f_1 f_2(a)$, $\mathrm{skel}_{\{f_1,f_2\}}(t) = f_2^2 f_1 f_2(a)$ and $\mathrm{height}(\mathrm{skel}_{\{f_1,f_2\}}(s)) = 3 < \mathrm{height}(\mathrm{skel}_{\{f_1,f_2\}}(t)) = 5$.

For further corollaries, we will need to consider the recursive path order on $T(\{f\} \cup \mathcal{F}_C)$ in more detail.

6.2. the case of one non-constant function symbol

In this section we let $\mathcal{F}, f$ and $\succ$ be as before and consider the recursive path order on $T(\mathcal{G})$ for $\mathcal{G} = \{f\} \cup \mathcal{F}_C$ in more detail. Since the case f is unary is not interesting as $(f^i(b) >_{\rm rpo} f^j(c)$ if and only if $i > j$ or $i = j$ and $b \succ c)$ we assume that f has arity at least 2.

We define the set of subterms at depth k of height j, $D_k^j(s)$, for a term s over an arbitrary set of symbols as follows.

Definition 6.7. Let $\mathcal{F}$ be a set of function symbols and $s \in T(\mathcal{F})$
if $s \in \mathcal{F}_C$, then

$$D_0^1(s) = \{s\}$$
$$D_i^j(s) = \emptyset \quad \text{otherwise}$$

if $s = g(s_1, \ldots, s_n)$, then

$$D_0^j(s) = \{s\} \qquad \text{if } j = \mathrm{height}(s)$$
$$D_0^j(s) = \emptyset \qquad \text{if } j \neq \mathrm{height}(s)$$
$$D_k^j(s) = \cup_{i=1}^n D_{k-1}^j(s_i) \qquad \text{otherwise.}$$

Note that if $D_k^j(s) \neq \emptyset$, then $0 \leq k \leq \mathrm{height}(s) - 1$, $1 \leq j \leq \mathrm{height}(s) - k$ and $j + k \leq \mathrm{height}(s)$.

Example 6.8. Let $s = g(g(a,a),a)$, then

$$\begin{aligned} D_1^2(s) &= D_0^2(g(a,a)) \cup D_0^2(a) \\ &= \{g(a,a)\} \\ D_1^1(s) &= D_0^1(g(a,a)) \cup D_0^1(a) \\ &= \{a\} \end{aligned}$$

and if $t = g(g(a,a), g(a,a))$, then $D_1^1(t) = D_0^1(g(a,a)) = \emptyset$.

Now we show that the recursive path order on $T(\mathcal{G})$ is an extension of an order obtained by first ordering by height and then comparing the terms of height p by comparing the largest subterms of height q at depth $p - q$ for some q.

Example 6.9. We will investigate the terms s, t from Example 6.1. We note that $\mathrm{height}(s) = \mathrm{height}(t)$. It is easy to check that $D_3^2(s) = \{f(a,a)\} = D_3^2(t)$, $D_2^3(s) = \{f(f(a,a), f(a,a))\}$ and $D_2^3(t) = \{f(f(a,a),a)\}$. It follows from the definition of the recursive path order that $f(f(a,a), f(a,a)) >_{\mathrm{rpo}} f(f(a,a),a)$. Thus, there is an element of $D_2^3(s)$ which dominates every element of $D_2^3(t)$. It then follows from the next theorem that $s >_{\mathrm{rpo}} t$.

Theorem 6.10. *Let $s, t \in T(\mathcal{F})$.*

(1) If $\mathrm{height}(s) > \mathrm{height}(t)$, then $s >_{\mathrm{rpo}} t$.
(2) Suppose $\mathrm{height}(s) = \mathrm{height}(t) = p$ and there is a q, $p \geq q \geq 1$ and an $s' \in D_{p-q}^q(s)$ such that $s' >_{\mathrm{rpo}} t'$ for each $t' \in D_{p-q}^q(t)$. Then $s >_{\mathrm{rpo}} t$.

Proof. (1) Is an instance of part 1 of Theorem 6.5, in which $s = \mathrm{skel}(s)$ and $t = \mathrm{skel}(t)$.
(2) The proof is by induction on $p - q$

Case $p - q = 0$. $s' \in D_0^p(s) = D_0^q(s) = \{s\}$ and $t' \in D_0^p(t) = D_0^q(t) = \{t\}$ so $s = s' >_{\mathrm{rpo}} t' = t$ as required.
Case $p > q$. Suppose $s = f(s_1, \ldots, s_n)$ and $t = f(t_1, \ldots, t_n)$. By hypothesis we have $s' \in D_{p-q}^q(s) = \cup_i D_{p-q-1}^q(s_i)$, so there is a j such that $s' \in D_{p-q-1}^q(s_j) \neq \emptyset$. It follows that $q + (p - q - 1) \leq \mathrm{height}(s_j) \leq p - 1$ and so $\mathrm{height}(s_j) = p - 1$. We now show that $s_j >_{\mathrm{rpo}} t_i$ for $i = 1, \ldots, n$. There are two cases:

Case $p - 1 = \mathrm{height}(s_j) > \mathrm{height}(t_i)$. Then $s_j >_{\mathrm{rpo}} t_i$ by part 1. of this theorem.

Case $p - 1 = \mathrm{height}(s_j) = \mathrm{height}(t_i)$. Now $p > q$, so $p - 1 \geq q \geq 1$ and $s' \in D_{p-q-1}^q(s_j)$, and if $t' \in D_{p-q-1}^q(t_i)$, then $t' \in D_{p-q}^q(t)$. It follows by the hypothesis of the theorem that $s' >_{\mathrm{rpo}} t'$. Now $(p-1) - q < p - q$ and so it follows from the induction hypothesis that $s_j >_{\mathrm{rpo}} t_i$.

Thus $s_j >_{\mathrm{rpo}} t_i$ for all i.

It follows that $s >_{\mathrm{rpo}} t$. □

We now give a simplified version of this in the case where $\mathcal{F}_C$ has one element, a. The unique maximal tree in f, a of height q is T_q defined by $T_1 = a$, $T_{r+1} = f(T_r, T_r)$. Any other tree of height q embeds in T_q and so if $\mathrm{height}(s) = \mathrm{height}(t) = p$ and for some value of q with $p \geq q \geq 1$ we find that $D^q_{p-q}(s)$ contains T_q and $D^q_{p-q}(t)$ does not, then $s >_{\rm rpo} t$. More formally:

Corollary 6.11. *Let* $s \in T(\mathcal{F})$, *with* $\mathrm{height}(s) = p$ *and let*

$$TD(s) = \max\{q | T_q \in D^q_{p-q}(s)\}.$$

Suppose that either $\mathrm{height}(s) > \mathrm{height}(t)$, *or* $\mathrm{height}(s) = \mathrm{height}(t)$ *and either* $TD(s) > TD(t)$. *Then* $s >_{\rm rpo} t$.

We note a phenomenon related to this work in the case of two unary function symbols, $\mathcal{F} = \{f, g, a\}$ with $f \succ g$. As we have seen in this case $\mathrm{height}(\mathrm{skel}(s)) = \#(f, s) + 1$. So for any s, t, $s >_{\rm rpo} t$ implies that $\#(f, s) \geq \#(f, t)$, and $\#(f, s) > \#(f, t)$ implies that $s >_{\rm rpo} t$, confirming what we already know, see Theorem 2.14.

If $\#(f, s) = \#(f, t) = i$ we may count occurrences of $\{u_j = f^j g f^{i-j} | j = 0, \ldots, i\}$ in s, t to obtain the following by a routine induction.

Theorem 6.12. *Let* $T = T(\{f, g, a\})$ *be a set of monadic terms and* $>_{\rm rpo}$ *be the recursive path order on* T *with* $f \succ g$. *Then* $s >_{\rm rpo} t$ *if and only if:*

(1) $\#(f, s) > \#(f, t)$ *or*
(2) $\#(f, s) = \#(f, t) = i$ *and* $(\#(u_0, s), \ldots, \#(u_i, s)) >_{\rm rpo}{}^{LexR} (\#(u_0, t), \ldots, \#(u_i, t))$.

7. Conclusions

We have extended our classification of simplification orders on monadic terms to orders on arbitrary terms T which induce a total simplification ordering on the spine of T. We associate to each such order $\succ$ on T numeric invariants which establish $\mathcal{P}_n$ as a classifying space for these orders, and hence associate to any rewrite system over T a subset of $\mathcal{P}_n$ corresponding to the orders which prove R terminating. The invariants induce a map $\bar{\mu} : T \to \mathbb{R}$, and hence give rise to a pre-order $\geq_{\bar{\mu}}$ on T which extends $\succ$. Note that our results do not require $\succ$ to be a simplification order: we have yet to explore how far our work is useful in analysing non-simplification orderings used in termination proofs, such as those of Zantema (1994).

One further step would be to refine the classification following the analysis of Martin and Scott (1997) and Prohle and Perlo-Freeman (1997) in the two letter case: another to incorporate our work on counting patterns in Section 5, so that we can calculate $>_{\bar{\mu}}$ directly from the orders $>_i$ on the patterns P_i. It seems apparent that our work on the recursive path order in Section 6 extends to other syntactic orders of this nature, and could provide a link with our main classification result. One might also consider matters such as order-types (Cichon and Weiermann, 1997) or decidability (Middeldorp and Gramlich, 1995) in our framework.

The eventual motivation for such work is an understanding of termination problems, and the establishment for terms of notions corresponding to the Gröbner walk and Gröbner fan of an ideal in a polynomial ring (Faugere *et al.*, 1993). This would enable us to understand in a uniform framework termination proofs for a rewrite system amenable to these methods.

Acknowledgements

This work was supported by the UK EPSRC under grant number GR/K75767, Geometry, Invariants and Termination. The authors also gratefully acknowledge the careful work of the referees.

References

Ben Cherifa, A., Lescanne, P. (1987). Termination of rewriting systems by polynomial interpretations and its implementation. *Sci. Comput. Program.*, **9**, 137–159.

Cichon, E. A., Weiermann, A. (1997). Term rewriting theory for the primitive recursive functions. *Ann. Pure Appl. Logic*, **83**, 199–223.

Cropper, N., Martin, U. (2000). The classification of polynomial orderings on monadic terms. *Appl. Algebra Eng., Commun. Comput.*

Dershowitz, N. (1982). Orderings for term-rewriting systems. *Theor. Comput. Sci.*, **17**, 279–301.

Dick, J., Kalmus, J., Martin, U. (1990). Automating the Knuth–Bendix ordering. *Acta Inform.*, **28**, 95–119.

Faugere, J., Gianni, P., Lazard, D., Mora, T. (1993). Efficient computation of zero-degree Gröbner bases by change of ordering. *J. Symb. Comput.*, **16**, 329–344.

Gallier, J. H. (1991). What's so special about Kruskal's theorem and the ordinal Γ_0? A survey of some results in proof theory. *Ann. Pure Appl. Logic*, **53**, 199–260.

Hofbauer, D. (1992). Termination proofs by multiset path ordering imply primitive recursive derivation lengths. *Theor. Comput. Sci.*, **105**, 129–140.

Knuth, D., Bendix, P. (1970). Simple word Problems in Universal Algebras. In Leech, J. ed., *Computational Problems in Abstract Algebra*, pp. 263–298. Oxford, Pergamon Press.

Lankford, D. S. (1975). Canonical algebraic simplification in computational logic. Technical Report ATP-25, University of Austin, Texas.

Lescanne, P. (1981). Decomposition ordering as a tool to prove the termination of rewriting systems. In *the Seventh IJCAI*, pp. 548–550.

Linton, S., Shand, D. (1996). Some group theoretic examples for completion theorem provers. *J. Autom. Reasoning.*, **17**, 145–169.

Martin, U. (1989). A geometrical approach to multiset orderings. *Theor. Comput. Sci.*, **67**, 37–54.

Martin, U. (1990). A note on division orderings on strings. *Inf. Process. Lett.*, **36**, 237–240.

Martin, U. (1993). Linear interpretations by counting patterns. In Kirchner, C. ed., *RTA '93*, LNCS **690**, pp. 421–433. New York, Springer.

Martin, U. (1995). On the diversity of orderings on strings. *Fundam. Inf.*, **24**, 25–46.

Martin, U. (1996). Theorem proving with group presentations: examples and questions. In McRobbie, M., Slaney, J. eds, *CADE-13*, LNCS **1104**, pp. 358–372. New York, Springer.

Martin, U., Scott, E. (1997). The order types of termination orderings on monadic terms, strings and multisets. *J. Symb. Logic*, **62**, 624–635.

Middeldorp, A., Gramlich, B. (1995). Simple termination is difficult. *Appl. Algebra Eng., Commun. Comput.*, **6**, 115–128.

Mora, T., Robbiano, L. (1986). The Gröbner fan of an ideal. *J. Symb. Comput.*, **6**, 183–208.

Prohle, P., Perlo-Freeman, S. (1997). Scott's conjecture is true, position sensitive weights. In Comon, H. ed., *RTA-97*, LNCS **1232**, pp. 217–228. New York, Springer.

Scott, E. (1994). Weights for total division orderings on strings. *Theor. Comput. Sci.*, **135**, 345–359.

Steinbach, J. (1995). Simplification orderings: history of results. *Fundam. Inf.*, **24**, 47–87.

Touzet, H. (1997). Propriétés combinatoires pour la terminaison de systèmes réécriture. Ph. D. Thesis, L'université Henri Poincaré - Nancy 1.

Zantema, H. (1994). Termination of term rewriting: interpretation and type elimination. *J. Symb. Comput.*, **17**, 23–50.

Originally Received 1 June 1998
Accepted 1 December 1999