A NOTE ON SIMPLIFICATION ORDERINGS *

Nachum DERSHOWITZ
Department of Computer Science, University of Illinois at Urbana-Champaign, Urbana, IL 61801, U.S.A.

Received 30 April 1979; revised version received 24 September 1979

Simplification orderings, term-rewriting, termination, well-founded orderings

1. A termination theorem

In this note we present a powerful method for proving the termination of term-rewriting systems based on the following notion of a simplification ordering.

Definition. A transitive and irreflexive relation $\succ$ is a *simplification ordering* on a set of terms T if for any terms $t, t', f(\cdots t \cdots), f(\cdots t' \cdots) \in T$

(1) $t \succ t'$ implies $f(\cdots t \cdots) \succ f(\cdots t' \cdots)$ and
(2) $f(\cdots t \cdots) \succ t$.

Note that this definition does not require that $\succ$ is well-founded (cf. [5,7]). We shall assume throughout this note that all function symbols have fixed arity.

A *term-rewriting system* **P** over a set of terms T is a finite set of rewrite rules of the form $\ell_i(\bar{\alpha}) \to r_i(\bar{\alpha})$, where the $\bar{\alpha}$ are variables ranging over T. Such a rule is applied to a term $t \in T$ in the following manner: if t contains a subterm $\ell_i(\bar{a})$, i.e. the variables $\bar{\alpha}$ are instantiated with terms $\bar{a}$, then replace that subterm with the corresponding term $r_i(\bar{a})$, thereby obtaining t'. The choice of rule and subterm is nondeterministic. We write $t \Rightarrow t'$ to indicate that the term t' can be derived from the term t by a single application of a rule in **P** to one of the subterms of t. (The variables appearing in r_i must be a subset of those in ℓ_i.) For example, the system consisting of the one rule $(\alpha \cdot \beta) \cdot \gamma \to \alpha \cdot (\beta \cdot \gamma)$ reparenthesizes a product by associating to the right. Applying that rule twice to the term $t = (a \cdot b) \cdot ((c \cdot d) \cdot e)$, we get

$$t \Rightarrow a \cdot (b \cdot ((c \cdot d) \cdot e)) \Rightarrow a \cdot (b \cdot (c \cdot (d \cdot e)))$$

or, alternatively,

$$t \Rightarrow (a \cdot b) \cdot (c \cdot (d \cdot e)) \Rightarrow a \cdot (b \cdot (c \cdot (d \cdot e)))\,.$$

In either case, no further applications of the rule are possible. We say that a term-rewriting system **P** *terminates*, if there exist no infinite sequences of terms $t_i \in T$ such that $t_1 \Rightarrow t_2 \Rightarrow t_3 \Rightarrow \cdots$. In general, it is undecidable whether a system terminates [2].

The following theorem gives a sufficient criterion for proving that a term-rewriting system terminates for all inputs.

Termination Theorem. A term-rewriting system $\mathbf{P} = \{\ell_i \to r_i\}_{i=1}^{p}$ over a set of terms T terminates if there exists a simplification ordering $\succ$ over T such that

$$\ell_i \succ r_i\,, \quad i = 1, ..., p,$$

for any assignment of terms in T to the variables of ℓ_i.

The proof of this theorem is based on the following:

Tree Theorem ([4]). In any infinite sequence $t_1, t_2, \ldots$ of terms over a finite set of function symbols, there exists a pair of terms t_i and t_j, $i < j$, such that t_i is homeomorphically embedded in t_j (when t_i and t_j are viewed as trees).

* This research was supported under NSF Grant MCS77-22830.

We shall denote this embedding relation by $\trianglelefteq$.
We have

$$s = f(s_1, s_2, ..., s_m) \trianglelefteq g(t_1, t_2, ..., t_n) = t\,,$$

if and only if

(a) $f = g$ and $s_i \trianglelefteq t_i$ for all i, $1 \leqslant i \leqslant m = n$ or else
(b) $s \trianglelefteq t_i$ for some i, $1 \leqslant i \leqslant n$.

Proof of Tree Theorem ([8]). Assume that the theorem is false. Then there exist one or more infinite 'counterexample' sequences of terms such that no element can be embedded in a subsequent one. We construct a 'minimal' counterexample sequence $t = (t_i)_i$ in the following manner: if the elements $t_1, t_2, ..., t_{i-1}$ $(i \geqslant 1)$ have already been chosen, then let t_i be a minimal size (i.e. number of function symbols) i^{th} element of a counterexample sequence beginning with the elements already chosen.

We need the following observation: Let $s = (s_i)_i$ be any infinite sequence whose elements are proper subterms of successive elements of some subsequence of t. By the assumption of minimality of t, if s_1 is a subterm of some t_k, then the sequence $t_1, t_2, ..., t_{k-1}, s_1, s_2, ...$ must contain an embedding pair $s_i \trianglelefteq s_j$ $(i < j)$. Moreover, there must exist an infinite chain $s_{i_1} \trianglelefteq s_{i_2} \trianglelefteq s_{i_3} \trianglelefteq \cdots$ of subterms $(i_1 < i_2 < i_3 < \cdots)$; each embedded in the next. (Otherwise, there would be an infinite number of maximal length, but finite, chains, each beginning beyond where the previous chain ended. But then the last elements of those chains form an infinite sequence of subterms which must also contain an embedding pair, implying that some chain was not maximal.)

Now, by the pigeon-hole principle, there exists an infinite subsequence $f(s_1^1, s_2^1, ..., s_m^1)$, $f(s_1^2, s_2^2, ..., s_m^2)$, ... of t, all elements of which have the same outermost function symbol, f. Furthermore, by the above observation, the sequence $(s_1^i)_i$ of first arguments contains an infinite embedding chain $s_1^{i_1} \trianglelefteq s_1^{i_2} \trianglelefteq \cdots$. Consider the sequence of terms $f(s_1^{i_1}, s_2^{i_1}, ..., s_m^{i_1})$, $f(s_1^{i_2}, s_2^{i_2}, ..., s_m^{i_2})$, ... Repeating the above procedure $m - 1$ times for the remaining arguments $s_2^i, ..., s_m^i$ yields an infinite subsequence of terms of t, every argument of which can be embedded in the corresponding argument of the subsequent term. Consequently, each term in the subsequence can be embedded in the next; hence, t could not have been a counterexample.

We shall need the following:

Lemma. Let s and t be terms in T. If $s \trianglelefteq t$, then $s \preccurlyeq t$ in any simplification ordering $\succ$ over T.

Proof. The proof is by induction on the size of t. Assume that $s' \trianglelefteq t'$ implies $s' \preccurlyeq t'$ for any t' smaller than t. By the definition of $\trianglelefteq$, if $s = f(s_1, s_2, ..., s_m) \trianglelefteq g(t_1, t_2, ..., t_n) = t$ (m or n may be 0), then either

(a) $f = g$ and $s_i \trianglelefteq t_i$ for all $1 \leqslant i \leqslant m = n$, in which case $s_i \preccurlyeq t_i$ and therefore $s \preccurlyeq t$ (property (1) of simplification orderings); or else
(b) $s \trianglelefteq t_i$ for some i, in which case $s \preccurlyeq t_i \prec g(\cdots t_i \cdots) = t$ (property (2)).

We are ready for the

Proof of Termination Theorem. Assume that **P** does not terminate. Then there exists an infinite sequence of terms $t_1 \Rightarrow t_2 \Rightarrow \cdots$. Note that there can be only a finite number of function symbols appearing in the sequence (those in t_1 and in **P**). If for a simplification ordering $\succ$ we have $\ell_i \succ r_i$, then it follows (using property (1)) that $t_1 \succ t_2 \succ \cdots$, and by transitivity that $t_i \succ t_j$ for all $i < j$. But, by the Tree Theorem $t_i \trianglelefteq t_j$ for some $i < j$, and by the lemma $t_i \preccurlyeq t_j$. This contradicts the asymmetry of $\succ$ (asymmetry follows from transitivity and irreflexivity).

This result may be used to simplify proofs of well-foundedness and termination, e.g. those in [1,3,6, 9,10].

2. Some examples

Consider the one-rule term-rewriting system

$$(\alpha \cdot \beta) \cdot \gamma \to \alpha \cdot (\beta \cdot \gamma)$$

and the following recursively defined ordering: $t \succ t'$ for two terms t and t', if

(i) $|t| > |t'|$, where $|t|$ denotes the number of function symbols in t, or else t and t' are products of the forms $\alpha \cdot \beta$ and $\alpha' \cdot \beta'$, respectively, and
(ii) $|t| = |t'|$ and $\alpha \succ \alpha'$ or else
(iii) $|t| = |t'|$, $\alpha = \alpha'$, and $\beta \succ \beta'$ (cf. the ordering in [3]).

To see that this is a simplification ordering, note that $t \succ t'$ implies $|t| \geq |t'|$. Thus, if $\alpha \succ \alpha'$, then $\alpha \cdot \beta \succ \alpha' \cdot \beta$ (by (i) or (ii)) and $\beta \cdot \alpha \succ \beta \cdot \alpha'$ (by (i) or (iii)). Furthermore, $|\alpha \cdot \beta| > |\alpha|, |\beta|$ and therefore $\alpha \cdot \beta \succ \alpha, \beta$ (by (i)). To prove termination, we need only verify that $(\alpha \cdot \beta) \cdot \gamma \succ \alpha \cdot (\beta \cdot \gamma)$ in this ordering. This follows (by (ii)) from the fact that $|(\alpha \cdot \beta) \cdot \gamma| = |\alpha \cdot (\beta \cdot \gamma)|$ and $|\alpha \cdot \beta| > |\alpha|$.

As a second example, we prove that the system

$$\alpha \cdot (\beta + \gamma) \to (\alpha \cdot \beta) + (\alpha \cdot \gamma)\,,$$

$$(\beta + \gamma) \cdot \alpha \to (\beta \cdot \alpha) + (\gamma \cdot \alpha)\,,$$

for distributing multiplication over addition, terminates. Let the function τ map terms into multisets recursively as follows:

(i) for any sum $\alpha + \beta$, $\tau(\alpha + \beta) = \tau(\alpha) \cup \tau(\beta)$, where $\cup$ denotes union of multisets,

(ii) for any product $\alpha \cdot \beta$, $\tau(\alpha \cdot \beta) = \{\tau(\alpha) \cup \tau(\beta)\}$, and

(iii) for any atomic term u, $\tau(u) = \{\emptyset\}$, where $\emptyset$ denotes the empty multiset.

For example, $\tau(((a + b) \cdot (c + d)) + e) = \{\{\emptyset, \emptyset, \emptyset, \emptyset\}, \emptyset\}$.

We use the following simplification ordering: $t \succ t'$, for two terms t and t', if $\tau(t) \gg \tau(t')$ in the *nested multiset ordering* $\gg$. In this ordering (see [1]), $X \cup Z \gg Y \cup Z$, for multisets $X \neq \emptyset$, Y, and Z, if for each element $y \in Y$ there is some $x \in X$ such that $x \gg y$. For example, $\{\{\emptyset, \emptyset, \emptyset, \emptyset\}, \emptyset\} \gg \{\{\emptyset, \emptyset, \emptyset\}, \{\emptyset, \emptyset, \emptyset\}, \emptyset\}$, since $\{\emptyset, \emptyset, \emptyset, \emptyset\} \gg \{\emptyset, \emptyset, \emptyset\}$. It should be obvious from this definition of $\gg$ that $X \cup Z \gg Y \cup Z$ if $X \gg Y$ and that $X \cup Z \gg Z$ if $X \neq \emptyset$. It is also easy to prove that $\{\cdots X \cdots\} \gg X$ for any multiset X, by induction on the nested depth of X. With these facts in mind, it is straightforward to verify that $\succ$ is a simplification ordering on terms.

To prove termination it remains to show that $\alpha \cdot (\beta + \gamma) \succ (\alpha \cdot \beta) + (\alpha \cdot \gamma)$ and $(\beta + \gamma) \cdot \alpha \succ (\beta \cdot \alpha) + (\gamma \cdot \alpha)$, i.e. we must show

$$\tau(\alpha \cdot (\beta + \gamma)) = \tau((\beta + \gamma) \cdot \alpha) = \{\tau(\alpha) \cup \tau(\beta) \cup \tau(\gamma)\}$$
$$\gg \tau((\alpha \cdot \beta) + (\alpha \cdot \gamma)) = \tau((\beta \cdot \alpha) + (\gamma \cdot \alpha))$$
$$= \{\tau(\alpha) \cup \tau(\beta), \tau(\alpha) \cup \tau(\gamma)\}\,.$$

This in turn follows from the fact that the multisets $\tau(\beta)$ and $\tau(\gamma)$ cannot be empty and therefore

$$\tau(\alpha) \cup \tau(\beta) \cup \tau(\gamma) \gg \tau(\alpha) \cup \tau(\beta), \tau(\alpha) \cup \tau(\gamma)\,.$$

3. An application

As an application of the Termination Theorem, consider the following class of orderings: with each function symbol f of arity $n \geq 0$ associate some real polynomial $F(x_1, ..., x_n)$ of n variables. Then extend this relation to a morphism on terms, i.e. $\overline{f(t_1, ..., t_n)} = F(\overline{t_1}, ..., \overline{t_n})$, where $\overline{t}$ is the real expression associated with the term t. The set of terms T constructed from those function symbols may be ordered according to the real values of the associated expressions, i.e. $t \succ t'$, for two terms t and t' in T, if and only if $\overline{t} > \overline{t'}$. Thus, to prove that a term-rewriting system **P** over T terminates, one must show that the polynomials $F(x_1, ..., x_n)$ satisfy the two conditions for simplification orderings

(1) $x > x'$ implies $F(\cdots x \cdots) > F(\cdots x' \cdots)$ and

(2) $F(\cdots x \cdots) > x$, and also that $\overline{\ell_i} > \overline{r_i}$ each rule $\ell_i \to r_i$ in **P**.

Conveniently, these are all decidable properties for polynomials over the reals [11]. By the same token, it is decidable if there exists any polynomial of degree less than any given n that demonstrates termination. In this manner, the undecidability for polynomials over the natural numbers, encountered in the method of [6], is circumvented.

Finally, we note that the Termination Theorem provides sufficient but not necessary conditions for termination. To see this, consider, for example, the one-rule system $ff\alpha \to fgf\alpha$, where f and g are unary function symbols. This system always terminates, since each application of the rule decreases the number of adjacent f's. On the other hand, $ff\alpha \trianglelefteq fgf\alpha$ and therefore $ff\alpha \preccurlyeq fgf\alpha$ in any simplification ordering. Consequently, there is no simplification ordering $\succ$ under which $ff\alpha \succ fgf\alpha$. Moreover, there can be no well-ordering $\succ$ of all the terms constructable from f and g – that satisfies the monotonicity property (1) – under which $ff\alpha \succ fgf\alpha$. (Since $\succ$ cannot be a simplification ordering, for some term α it must be that $\alpha \succ h\alpha$, where h is f or g; consequently, $\alpha \succ h\alpha \succ hh\alpha \succ \cdots$ is an infinite descending sequence and $\succ$ cannot be a well-ordering.)

References

[1] N. Dershowitz and Z. Manna, Proving termination with multiset orderings Comm. ACM 22 (8) (1979) 465–476.

[2] G. Huet and D.S. Lankford, On the uniform halting problem for term rewriting systems, Report 283, IRIA, Le Chesney, France (1978).

[3] D.E. Knuth and P.B. Bendix, Simple word problems in universal algebras, in: J. Leech, Ed., Computational Problems in Universal Algebras (Pergamon Press, Oxford, 1969) 263–297.

[4] J.B. Kruskal, Well-quasi-ordering, the tree theorem, and Vazsonyi's conjecture, Trans. AMS 95 (1960) 210–225.

[5] D.S. Lankford, Canonical algebraic simplification in computational logic, Memo ATP-25, Automatic Theorem Proving Project, Univ. of Texas, Austin, TX (1975).

[6] D.S. Lankford, On proving term rewriting systems are Noetherian, Memo MTP-3, Mathematics Dept., Louisiana Tech. Univ., Ruston, LA (1979).

[7] Z. Manna and S. Ness, On the termination of Markov algorithms, in: Proc. Third Hawaii Internat. Conf. on System Sciences, Honolulu, HI (1970) 789–792.

[8] C.St.J.A. Nash-Williams, On well-quasi-ordering finite trees, Proc. Cambridge Philos. Soc. 59 (1963) 833–835.

[9] D. Plaisted, Well-founded orderings for proving the termination of rewrite rules, Report R-78-932, Dept. of Computer Science, Univ. of Illinois, Urbana, IL (1978).

[10] D. Plaisted, A recursively defined ordering for proving termination of term rewriting systems, Report R-78-943, Dept. of Computer Science, Univ. of Illinois, Urbana, IL (1978).

[11] A. Tarski, A Decision Method for Elementary Algebra and Geometry (Univ. of California Press, Berkeley, CA, 1951).