J. Symbolic Computation (1997) **23**, 399–411

Omega-Termination is Undecidable for Totally Terminating Term Rewriting Systems

ALFONS GESER†

Wilhelm-Schickard-Institut für Informatik, Universität Tübingen, Sand 13, D-72076 Tübingen, Germany

(Received 11 April 1996)

We give a complete proof of the fact that the following problem is undecidable:

Given: A term rewriting system, where the termination of its rewrite relation is provable by a total reduction order on ground terms,
Wanted: Does there exist a strictly monotonic interpretation in the positive integers that proves termination?

1. Introduction

Termination of a term rewriting system (TRS), i.e. the non-existence of an infinite rewrite reduction, $t_1 \to_R t_2 \to_R \cdots$, is one of the key notions in term rewriting. It is the basis of a number of decision algorithms for properties undecidable in the general case. In this paper we assume that TRSs have finitely many rewrite rules.

The termination of a TRS R means precisely the existence of a well-founded order $>$ on terms, closed under contexts and substitution—a *reduction order*—such that every rule $l \to r$ in R is ordered $l > r$ (Manna and Ness, 1970). Such an order can always be derived from a well-founded order on ground terms that is closed under contexts, whence one may restrict attention to such orders.

Terminating TRSs may now be further distinguished by some additional properties $>$ may have. If $>$ does not order any $s > t$ where s is homeomorphically embedded in t then R is called *non-self-embedding*. If $>$ orders subterms, i.e. $s > t$ holds if t is a proper subterm of s, then R is called *simply terminating*. If $>$ is *total on ground terms*, i.e. $s > t$ or $t > s$ holds for all ground terms $s \neq t$, then R is called *totally terminating*. Equivalently, R is totally terminating if $>$ is the reduction order induced by a strictly monotonic interpretation into the ordinal numbers. The smallest ordinal section for which this is possible is then also called the *termination type* of R. If $>$ is induced by a strictly monotonic interpretation into the positive integers (the well-ordered set ω) then R is called ω-*terminating*.

† E-mail: geser@informatik.uni-tuebingen.de

0747–7171/97/040399 + 13 $25.00/0 sy960095

These specializations form a true hierarchy, called the *termination hierarchy* (Zantema, 1994).

termination $\supsetneq$ non-self-embedding $\supsetneq$ simple termination $\supsetneq$ total termination $\supsetneq$ ω-termination $\supsetneq$ polynomial termination

There are two reasons to study the properties stronger than termination. First, they obey better decomposition theorems. For instance, ω-termination and simple termination satisfy direct sum modularity (Kurihara and Ohuchi, 1990), and total termination allows distribution elimination without linearity conditions (Zantema, 1994). Termination itself does not have these properties.

Second, there are tools that provide proofs for these properties. For instance, strictly monotonic interpretations into the positive integers prove ω-termination, and so does the recursive path order (Meeussen and Zantema, 1993) for finite TRSs. The Knuth/Bendix order, and strictly monotonic interpretations into the ordinal numbers each prove total termination (Ferreira, 1995).

It is known that termination is undecidable for TRSs, even for the restricted case of TRS with only unary function symbols (Huet and Lankford, 1978), and for the case of one-rule TRSs (Dauchet, 1992).

Each of the upper four properties is known to be undecidable (Huet and Lankford, 1978; Plaisted, 1985; Caron, 1991; Zantema, 1995). Huet and Lankford (1978) proved undecidability of termination by reducing the halting problem of Turing machines to it. Lescanne (1994) proposed to use the technically simpler reduction from the Post Correspondence Problem.

The present paper addresses the next question: Whether undecidability still holds for ω-termination. We study an even more difficult problem:

Given: A totally terminating TRS R.
Wanted: Is R ω-terminating?

The question can be rephrased as: "Has R termination type ω?". We prove by a reduction from the Post Correspondence Problem that this problem is *undecidable.* As a consequence, ω-termination is an undecidable property of TRSs.

2. Basic Notions

We assume that the reader is familiar with termination of TRSs (Dershowitz, 1987). We use notation as summarized in Dershowitz and Jouannaud (1991).

We consider a fixed signature, and may therefore write $\mathcal{GT}$ for the set of ground terms, i.e. terms that do not contain variables.

Post's well-known Correspondence Problem (PCP) is defined as follows. Given a finite alphabet Γ, a PCP *instance* is a finite binary relation $P \subseteq \Gamma^+ \times \Gamma^+$ on proper words over Γ. The PCP instance P is said to *have a solution* $\gamma \in \Gamma^+$ if

$$\alpha_1\alpha_2 \cdots \alpha_n = \gamma = \beta_1\beta_2 \cdots \beta_n$$

holds where $(\alpha_i, \beta_i) \in P$ for all $1 \leq i \leq n$. If Γ has at least two elements then the problem: given P, has P a solution? is undecidable (Post, 1946). Henceforth let $\Gamma = \{0, 1\}$.

3. Plan

We start off from Zantema's standard example of a totally terminating, not ω-terminating TRS.

THEOREM 3.1. *The one-rule TRS*

$$h(g(x)) \to g(h(h(x))) \tag{3.1}$$

is totally terminating but not ω-terminating.

PROOF. ((ZANTEMA, 1994)) To show total termination one gives the strictly monotonic interpretation $[h](m,n) = (m, m+n)$, $[g](m,n) = (2m+1, n)$ in pairs of positive integer numbers, lexicographically ordered. This interpretation is strictly monotonic. Moreover the rule is ordered by the induced interpretation order:

$$\begin{aligned}[h(g(t))] = [h]([g](m,n)) &= [h](2m+1,n)\\ &= (2m+1, 2m+1+n) >_{lex} (2m+1, 2m+n)\\ &= [g](m,2m+n) = [g]([h](m,m+n)) = [g]([h]([h](m,n))) = [g(h(h(t)))]\end{aligned}$$

holds for all $t \in \mathcal{GT}$ where $(m,n) =_{\text{def}} [t]$.

To disprove ω-termination one first observes that every strictly monotonic interpretation into the positive integers entails that the induced reduction order is finitely branching. We claim that $[g]([t]) > [h]^k([t])$ holds for all $k \in \mathbb{N}$ and $t \in \mathcal{GT}$.

Proof by induction on k: The base case, $[g]([t]) > [t] = [h]^0([t])$, follows by the fact that a total reduction order must order subterms. The inductive case is solved through

$$[h]([g]([t])) > [g]([h]([h]([t]))) > [h]^{k-1}([h]([h]([t]))) = [h]^{k+1}([t]),$$

by the fact that the rule is ordered, and by the inductive hypothesis, respectively. The claim, $[g]([t]) > [h]^k([t])$, follows by the cancellation rule for total reduction orders.

By the property just proven, $[g]([t])$ is an upper bound to infinitely many positive integers, $[h]^k([t])$, $k \in \mathbb{N}$, a contradiction. $\square$

We will use a PCP instance, P, to switch between this system, and another system which is ω-terminating,

$$h(g(x)) \to g'(h(h(x)))\,. \tag{3.2}$$

To prove ω-termination of this system, let $[g](x) = x+3$, $[g'](x) = [h](x) = x+1$ be an interpretation into the positive integers, naturally ordered. This interpretation is strictly monotonic and

$$[h(g(t))] = [h]([g]([t])) = [t]+4 > [t]+3 = [g']([h]([h]([t]))) = [g'(h(h(t)))]$$

holds for all $t \in \mathcal{GT}$. So system (3.2) is indeed ω-terminating.

If P has a solution, then the resulting system will have a behaviour comparable to system (3.1), else to system (3.2).

However, this is a strongly simplified picture; the detailed technical treatment of this idea is much more involved. Now we are going to put these ideas in a precise form.

To every letter $a \in \Gamma$ let there be a unique new barred letter $\bar{a}$. By small Greek letters we denote strings over Γ. For $\alpha = a_1 a_2 \ldots a_n$ let $\bar{\alpha} =_{\text{def}} \bar{a_n} \bar{a_{n-1}} \ldots \bar{a_1}$, the string of letters of α barred and *in reversed order.* We will consider barred and unbarred letters also as unary function symbols, and strings of barred and unbarred letters as contexts. Apart from that we fix another unary function symbol h, two 5-ary function symbols f and g, and a constant c.

Definition 3.1. *To every PCP instance P let a TRS R_P be assigned that contains exactly the rules*

$$h(g(\alpha(x), c, \beta(z), c, u)) \to f(x, \bar{\alpha}(c), z, \bar{\beta}(c), u) \tag{3.3}$$

$$f(\alpha(x), y, \beta(z), w, u) \to f(x, \bar{\alpha}(y), z, \bar{\beta}(w), u) \tag{3.4}$$

for each $(\alpha, \beta) \in P$, and the rules

$$f(c, \bar{a}(y), c, \bar{a}(w), u) \to g(a(c), y, a(c), w, h(h(u))) \tag{3.5}$$

$$g(x, \bar{a}(y), z, \bar{a}(w), u) \to g(a(x), y, a(z), w, u) \tag{3.6}$$

for each $a \in \Gamma$.

R_P is finite; it contains $2|P| + 2|\Gamma|$ rules. The following claims will be proven each in one successive section.

Theorem 3.2. *R_P is totally terminating.*

Theorem 3.3. *If the PCP instance P has a solution then R_P is not ω-terminating.*

Theorem 3.4. *If the PCP instance P has no solution then R_P is ω-terminating.*

These three statements together suffice to establish a proof of our claim.

Corollary. The following problem is undecidable.

> **Given:** A totally terminating TRS R.
> **Wanted:** Is R ω-terminating?

Proof. If the problem were decidable, then we could particularly employ the putative procedure for TRSs of the form R_P where P is a PCP instance, thanks to Theorem 3.2. But then we could use the procedure via Theorem 3.3 and Theorem 3.4 to solve the problem whether, given a PCP instance P, it has a solution. Contradiction to the well-known undecidability of the PCP (Post, 1946). □

4. Total Termination

Usually people prove total termination by a strictly monotonic interpretation in the positive integers, or by a lexicographic recursive path order (RPO). We surely cannot use the first method as we intend to prove that under certain circumstances (the PCP instance P has a solution) such an interpretation does not exist. For the same reason we cannot use an RPO either.

To prove that R_P is totally terminating, no matter whether the PCP instance P has a solution or not, we have to employ a method that establishes total termination even in the non-ω-termination case. The well-known Knuth/Bendix order is such a method.

DEFINITION 4.1. (KNUTH/BENDIX ORDER (KNUTH AND BENDIX, 1970)) *Let* $\gtrsim_{prec}$ *be a well-founded quasiorder on* $\mathcal{F}$*, and* $[\_]$ *be a linear monotonic interpretation in the positive integers such that* $[f](x_1, \ldots, x_n) = x_i$ *has a solution for some* i *only if* $n = 1$ *and* $f \gtrsim g$ *holds for every function symbol* $g \in \mathcal{F}$*. Then* $\gtrsim_{kbo} \subseteq \mathcal{GT} \times \mathcal{GT}$ *is defined by* $s = f(s_1, \ldots, s_m) \gtrsim_{kbo} g(t_1, \ldots, t_n) = t$ *if*

$$[s] > [t], \text{ or}$$
$$[s] = [t], \text{ and } f >_{prec} g, \text{ or}$$
$$[s] = [t], \text{ and } f \sim_{prec} g, \text{ and } (s_{\pi(1)}, \ldots, s_{\pi(m)}) \gtrsim_{kbo,lex} (t_{\rho(1)}, \ldots, t_{\rho(n)}),$$

where f *and* g *have lexicographic status* π *and* ρ *respectively.*

THEOREM 4.1. ((KNUTH AND BENDIX, 1970; DICK *et al.*, 1990)) *Under the conditions of Definition 4.1, the Knuth/Bendix order* $>_{kbo}$ *is a reduction order.*

The Knuth/Bendix order is also suitable to prove total termination, by a result of Ferreira.

THEOREM 4.2. ((FERREIRA, 1995), THEOREM 4.47) *Under the conditions of Definition 4.1, the Knuth/Bendix order* $>_{kbo}$ *can be extended to a reduction order that is total on ground terms. Hence any term rewriting system* R *that satisfies* $l\sigma >_{kbo} r\sigma$ *for any ground instance* $l\sigma \to r\sigma$ *of a rule* $l \to r$ *in* R *is totally terminating.*

In our setting we choose the interpretation ("weight function") $[\_]$ in $\mathbb{N}$ defined by

$$\begin{aligned}
&[c] = 1\\
&[0](x) = [1](x) = [\bar{0}](x) = [\bar{1}](x) = x + 1\\
&[f](x, y, z, w, u) = [g](x, y, z, w, u) = x + y + x + w + u + 1\\
&[h](x) = x
\end{aligned}$$

where the results are ordered by $>$ on the positive integers, and the precedence $h >_{prec} f >_{prec} g >_{prec} 1 >_{prec} 0 >_{prec} \bar{1} >_{prec} \bar{0} >_{prec} c$, and status f left-to-right, and g right-to-left. With that, the Knuth/Bendix order orders each rule. By Theorem 4.2 it follows that R_P is totally terminating.

5. If P has a Solution

We show in this section that in case the PCP instance P has a solution, $\gamma \in \Gamma^+$ say, then the TRS R_P is not ω-terminating.

Let P have a solution γ. Then we have for all ground terms t a R_P-reduction

$$\begin{aligned}
h(g(\gamma(c), c, \gamma(c), c, t)) &\to_{(3.3)} \to^*_{(3.4)} f(c, \bar{\gamma}(c), c, \bar{\gamma}(c), t)\\
&\to_{(3.5)} \to^*_{(3.6)} g(\gamma(c), c, \gamma(c), c, h(h(t)))\,. \qquad (5.1)
\end{aligned}$$

For the sake of contradiction assume now that there is a strictly monotonic interpretation $[h] : \mathbb{N}_+ \to \mathbb{N}_+$ and $[g] : \mathbb{N}_+^5 \to \mathbb{N}_+$ in the positive integer numbers, ordered by $>$ that orders R_P.

LEMMA 5.1. *If* $\gamma \in \Gamma^+$ *is a solution on the PCP instance* P*, then for all* $k \in \mathbb{N}$ *and* $t \in \mathcal{GT}$*, we have* $[g(\gamma(c), c, \gamma(c), c, t)] > [h^k(t)]$.

Proof. The following proof is the same in spirit as the one of Theorem 3.1.

Let $x = [\gamma(c)] = z$ and $y = [c] = w$ be fixed. We prove that for all $u =_{\mathrm{def}} [t]$ we have $[g](x, y, z, w, u) > [h]^k(u)$, by induction on k. Since the order $>$ is total, $>$ orders subterms. So $[g](x, y, z, w, u) > u$ which solves the case $k = 0$. Otherwise,

$$\begin{aligned} [h]([g](x, y, z, w, u)) &> [g](x, y, z, w, [h]([h](u))) && (*) \\ &> [h]^{k-1}([h]([h](u))) && \text{(ind. hyp.)} \\ &= [h]([h]^{k+1}(u)) \end{aligned}$$

where the step labelled by $(*)$ is justified because (5.1) holds, and $>$ is a reduction order that orders every ground instance of a rule in R_P. The claim follows by the cancellation rule for total reduction orders. □

By Lemma 5.1 $[g(\gamma(c), c, \gamma(c), c, t)]$ is an upper bound to infinitely many numbers $[t] < [h(t)] < [h^2(t)] < \cdots$. This is impossible in $\mathbb{N}_+$.

So, if P has a solution, the assumption that R_P were ω-terminating, is false.

6. If P has No Solution

In this section finally, we are going to show that in the case P has no solution, R_P indeed can be ordered by a strictly monotonic interpretation in the positive integers.

For convenience let us get rid of h symbols. To this end we define the term rewriting system R'_P by $R'_P = R_P \setminus \{(3.3), (3.5)\} \cup \{(6.1), (6.2)\}$ where Rule (6.1) and (6.2) are defined by

$$g(\alpha(x), c, \beta(z), c, u) \to f(x, \bar{\alpha}(c), z, \bar{\beta}(c), u), \tag{6.1}$$

$$f(c, \bar{a}(y), c, \bar{a}(w), u) \to g(a(c), y, a(c), w, u)\,. \tag{6.2}$$

By deletion of h symbols, every step $h^k(\Phi(p, q, r, s, t)) \to_{R_P} h^{k'}(\Phi'(p', q', r', s', t'))$ is mapped to a step $\Phi(p, q, r, s, t) \to_{R'_P} \Phi'(p', q', r', s', t)$.

Henceforth let P be a PCP instance that has no solution.

First we derive upper bounds for the length of R'_P, and so of R_P, reductions (Subsection 6.1). We then encode the information about maximal reduction lengths in an interpretation (Subsection 6.2) of which we finally prove that it orders R_P (Subsection 6.3) and is strictly monotonic (Subsection 6.4). It follows that R_P is ω-terminating.

6.1. upper bounds for reduction lengths

In case P has no solution one gets an upper bound for reduction lengths which can then be used to encode part of the interpretation.

For a ground term t, let $|t|$ denote the number of barred or unbarred letters in t not below an f, g, or h symbol.

Lemma 6.1. *Let P be a PCP instance that has no solution. Then*

1. *No ground term $g(p, q, r, s, t)$ starts a R'_P-reduction with more than*

$$\min\{|q|, |s|\} + 2 \cdot \min\{|p| + |q|, |r| + |s|\}$$

steps at the outermost f or g position.

2. *No ground term $f(p,q,r,s,t)$ starts a R'_P-reduction with more than*

$$\min\{|p|,|r|\} + 2\cdot\min\{|p|+|q|,|r|+|s|\}$$

steps at the outermost f or g position.

PROOF. We show only the proof of claim 1 below; claim 2 is proven in the same spirit.

Assume for the sake of contradiction that there is a R'_P-reduction starting from $g(p,q,r,s,t)$ which is longer. We are going to extract a solution of P from this reduction, a contradiction to the premise.

By the form of the rules, the reduction must be a prefix of some reduction

$$\begin{aligned}
g(p,q,r,s,t) &\to^*_{P1} g(\gamma(p),q',\gamma(r),s',t)\\
&\to_{(6.1)} f(p',\bar{\alpha_1}(c),r',\bar{\beta}_1(c),t)\\
&\to^*_{P2} f(p'',\bar{\alpha_i}\bar{\alpha_{i-1}}\cdots\bar{\alpha_1}(c),r'',\bar{\beta}_i\bar{\beta}_{i-1}\cdots\bar{\beta}_1(c),t)\\
&\to_{(6.2)} g(\alpha_i(c),\bar{\alpha_{i-1}}\cdots\bar{\alpha_1}(c),\beta_1(c),\bar{\beta}_{i-1}\cdots\bar{\beta}_1(c),t)\\
&\to^*_{P3} g(\delta(c),q'',\delta(c),s'',t)\\
&\to_{(6.1)} f(\dots)\\
&\to \cdots
\end{aligned}$$

where during $\to^*_{P1}$ every term has top symbol g, during $\to^*_{P2}$ every term has top symbol f, during $\to^*_{P3}$ every term has top symbol g.

We distinguish the phases $P1, P2, P3$ of which we deduce upper bounds of their reduction length.

During $P1$ only Rule (3.6) can be applied at the top. Therefore $P1$ has at most $\min\{|q|,|s|\}$ top steps, since at each top step the length of the first and third argument of the outermost g is decreased each by one.

In order to arrive at $P2$, one must have $q' = c = s'$ and so $q = \bar{\gamma}(c) = s$. During $P2$ only Rule (3.4) is applicable at the top; at most $\min\{|p|+|q|,|r|+|s|\} - 1$ times since at each top (3.4) or (6.1) step at least one letter of $\gamma(p)$ and of $\gamma(r)$ is consumed.

Only if $p'' = c = r''$, and so

$$\begin{aligned}
\gamma(p) &= \alpha_1\alpha_2\cdots\alpha_i(c)\\
\gamma(r) &= \beta_1\beta_2\cdots\beta_i(c)
\end{aligned}$$

does one arrive at $P3$, provided that $(\alpha_1,\beta_1),(\alpha_2,\beta_2),\dots,(\alpha_i,\beta_i) \in P$. Phase $P3$ takes at most $\min\{|p|+|q|,|r|+|s|\} - 1$ steps again, since only Rule (3.6) is applicable at the top, which each time consumes one letter of $\gamma(p)$ and one of $\gamma(r)$.

If our given reduction has length at least $\min\{|q|,|s|\} + 2\cdot\min\{|p|+|q|,|r|+|s|\} + 1$, then it must have passed the step after $P3$. But then $q'' = c = s''$ and so

$$\alpha_1\alpha_2\cdots\alpha_i = \delta = \beta_1\beta_2\cdots\beta_i$$

is a solution of the PCP instance P. □

With the aim of formalizing the notion of progress, we now define mutually recursively two functions, len_f and len_g, from quadruples of ground terms to non-negative integers.

$$\begin{aligned}
&\mathsf{len}_f(p,q,r,s) = \\
&\quad \max\{\mathsf{len}_g(a(c),q',a(c),s')+1 \mid p=c=r \wedge q=\bar{a}(q') \wedge s=\bar{a}(s') \wedge a\in\Gamma\} \cup \\
&\qquad \{\mathsf{len}_f(p',\bar{\alpha}(q),r',\bar{\beta}(s))+1 \mid p=\alpha(p') \wedge r=\beta(r') \wedge (\alpha,\beta)\in P\} \\
&\mathsf{len}_g(p,q,r,s) = \\
&\quad \max\{\mathsf{len}_f(p',\bar{\alpha}(c),r',\bar{\beta}(c))+1 \mid q=c=s \wedge p=\alpha(p') \wedge r=\beta(r') \wedge (\alpha,\beta)\in P\} \cup \\
&\qquad \{\mathsf{len}_g(a(p),q',a(r),s')+1 \mid q=\bar{a}(q') \wedge s=\bar{a}(s') \wedge a\in\Gamma\}
\end{aligned}$$

The recursive definition follows the structure of R'_P rules; if $\Phi,\Phi'\in\{f,g\}$ then $\mathsf{len}_\Phi(p,q,r,s)$ is defined as the maximum of all $\mathsf{len}_{\Phi'}(p',q',r',s')+1$ such that $\Phi(p,q,r,s,t)\to_{R'_P}\Phi'(p',q',r',s',t')$. Here we take the view that $\max\emptyset=0$ whence for instance $\mathsf{len}_f(c,c,c,c)=0$ holds.

The call $\mathsf{len}_\Phi(p,q,r,s)$ therefore computes the maximum length, provided it exists, of R'_P reductions starting from a ground term $\Phi(p,q,r,s,t)$. Termination of the recursive definition, and hence totality of len_f and len_g in the case where P has no solution is ensured by Lemma 6.1. The following are immediate consequences that we will use later.

Proposition 6.1. *Let P be a PCP instance that has no solution. Then for all ground terms p,q,r,s, and function symbols $\Phi\in\{f,g\}$,*

$$\mathsf{len}_\Phi(p,q,r,s) \le 3\cdot\min\{|p|+|q|,|r|+|s|\}\,.$$

Proposition 6.2. *Let P be a PCP instance that has no solution. Then for every ground instance $\Phi(p,q,r,s,t)\to\Phi'(p',q',r',s',t')$ of a rule in R'_P one gets $\mathsf{len}_\Phi(p,q,r,s) > \mathsf{len}_{\Phi'}(p',q',r',s')$.*

6.2. an interpretation

Now we are going to extract an interpretation in the positive integers from this knowledge. To this end we will utilize the fact that a positive integer can also be regarded as a sequence of decimal digits. The decimal system is only preferred for its familiarity.

First let us define a few useful auxiliary functions on $\mathbb{N}_+$. Let $\ell(z)$ denote the number of decimal digits of the positive integer number z.

$$\begin{gathered}
x\circ y =_{\text{def}} 10^{\ell(y)}\cdot x+y, \\
x^{\circ\,1} =_{\text{def}} x, \qquad x^{\circ\,y+1} =_{\text{def}} x^{\circ\,y}\circ x, \\
\mathsf{revc}(x,y) =_{\text{def}} x\circ y\circ 4^{\circ\,\ell(x)}, \\
\mathsf{bound}(x,y,z,w) =_{\text{def}} 3\cdot\min\{\ell(\mathsf{revc}(x,y)),\ell(\mathsf{revc}(z,w))\}-2\,.
\end{gathered}$$

Let $\circ$ bind weaker than $+$ or $\cdot$. Informally, $x\circ y$ yields the concatenation of the digit sequences of x and y without their leading zeros. This function is obviously associative whence one may drop parentheses. The expression $x^{\circ\,y}$ denotes the y-fold repetition of x.

A function $\pi:\mathbb{N}_+\to\mathcal{GT}$ from positive integers to ground terms is defined recursively as follows.

$$\pi(x' \circ 2) = 0(\pi(x')), \qquad \pi(2 \circ x' \circ 4) = \bar{0}(\pi(x')), \qquad \pi(1) = c,$$
$$\pi(x' \circ 3) = 1(\pi(x')), \qquad \pi(3 \circ x' \circ 4) = \bar{1}(\pi(x')), \qquad \pi(x) = h(c) \text{ else}$$

With that the definitions of the interpretation functions read as follows.

$$[c] = 1 \qquad [0](x) = x \circ 2 \qquad [1](x) = x \circ 3$$
$$[h](x) = x \circ 5 \qquad [\bar{0}](x) = 2 \circ x \circ 4 \qquad [\bar{1}](x) = 3 \circ x \circ 4$$
$$[f](x, y, z, w, u) = (u \cdot 156^{(\mathsf{len}_f(\pi(x),\pi(y),\pi(z),\pi(w))+1+B)}) \circ 6$$
$$[g](x, y, z, w, u) = (u \cdot 156^{(\mathsf{len}_g(\pi(x),\pi(y),\pi(z),\pi(w))+1+B)}) \circ 7$$

where B abbreviates for $B =_{\text{def}} (\mathsf{bound}(x, y, z, w) + 1)(\mathsf{revc}(x, y) + \mathsf{revc}(z, w))$.

A few words of explanation and a plan of the proof are in order. The complexity of the construction arises from the necessity to express functions on ground terms by (strictly) monotonic functions on positive integers being or not being interpretations of ground terms.

The interpretation mapping of a ground term t to the positive integer number $[t]$ is designed such that essential information about t can be retrieved from $[t]$. More specifically, the last digit of $[t]$ determines the top function symbol of t, unless that digit is 4, in which case the first digit, too, is needed to distinguish between $\bar{0}$ and $\bar{1}$.

The function $\pi : \mathbb{N}_+ \to \mathcal{GT}$ is designed to exploit this phenomenon to a certain degree. It retrieves enough information to tell the function len a parameter tuple such that len yields the correct maximal reduction length from a tuple of interpretations.

As on the one hand, $[\_]$ is not surjective on $\mathbb{N}_+$, and on the other hand, we are only interested in the leading string of $0, 1, \bar{0}, \bar{1}$, and potentially c, we have π map all numbers not ending in $0, \ldots, 4$, or ending in 4 but not beginning with 2 or 3, to $h(c)$, a representative of all terms that fail to match any symbol of R_P.

The function $(x, y, z, w) \mapsto \mathsf{len}_f(\pi(x), \pi(y), \pi(z), \pi(w))$ strictly decreases with each R_P rule, is however not monotonic. Therefore we override its growth by a function that is strictly monotonic, revc, stretched by a factor, bound, that is monotonic and an upper bound of len_f (Lemma 6.5). Both are invariant by R_P steps (Lemma 6.3) and so do not interfere the task to order R_P.

6.3. THE INTERPRETATION ORDERS R_P

For the proof that rules are ordered, we need a lemma saying that len_Φ is robust w.r.t. $\pi([\_])$ and a lemma saying that revc is robust against shovelling of strings.

LEMMA 6.2. *For all ground terms p, q, r, s, and function symbols $\Phi \in \{f, g\}$,*

$$\mathsf{len}_\Phi(p, q, r, s) = \mathsf{len}_\Phi(\pi([p]), \pi([q]), \pi([r]), \pi([s]))\,.$$

PROOF. Structural induction on ζ shows that for all strings $\zeta \in (\Gamma \cup \bar{\Gamma})^*$ of barred or unbarred letters, and for all ground terms t,

$$\pi([t]) = \zeta(t') \quad \Longleftrightarrow \quad \exists t''.\, t = \zeta(t'') \wedge \pi([t'']) = t'\,.$$

As an immediate consequence one gets that the first four arguments of $\Phi, \Phi' \in \{f, g\}$ are not sensitive to $\pi([\_])$. More precisely, for all ground terms $p, q, r, s, t, p', q', r', s', t'$,

$$\Phi(p, q, r, s, t) \to_{R'_P} \Phi'(p', q', r', s', t')$$

if and only if

$$\Phi(\pi([p]), \pi([q]), \pi([r]), \pi([s]), t) \to_{R'_P} \Phi'(\pi([p']), \pi([q']), \pi([r']), \pi([s']), t)\,.$$

The claim follows by definition of len_Φ. □

The strange exception of the interpretation of barred symbols is motivated to achieve the following effect.

EXAMPLE 6.1. *We have* $[01c] > [10c]$ *by*

$$\begin{aligned}[01c] &= (1 \circ 3) \circ 2 = 132 \\ [10c] &= (1 \circ 2) \circ 3 = 123\end{aligned}$$

but $[\bar{0}1c] = [\bar{1}\bar{0}c] > [\bar{0}\bar{1}c] = [\bar{1}0c]$ *by*

$$\begin{aligned}[\bar{1}\bar{0}c] &= 3 \circ (2 \circ 1 \circ 4) \circ 4 = 32144 \\ [\bar{0}\bar{1}c] &= 2 \circ (3 \circ 1 \circ 4) \circ 4 = 23144\,.\end{aligned}$$

The reader will easily find out that the crucial sequence of digits to compare is the same. Unbarred sequences are compared by their digits right-to-left, barred sequences left-to-right. This behaviour is intended, for we aim at the following general result.

LEMMA 6.3. *For all* $\alpha \in \Gamma^*$,

$$\mathsf{revc}([\alpha](x), y) = \mathsf{revc}(x, [\bar{\alpha}](y))\,.$$

PROOF. By structural induction on α. If α is the empty string, then the claim is trivial. Else let $\alpha = a\alpha'$ and let $d = 2$ if $a = 0$ and $d = 3$ if $a = 1$. Then

$$\begin{aligned}\mathsf{revc}([a]([\alpha'](x)), y) &= \mathsf{revc}([\alpha'](x) \circ d, y) && \text{(defn. } [a]) \\ &= \mathsf{revc}([\alpha'](x), d \circ y \circ 4) && \text{(defn. } \mathsf{revc}) \\ &= \mathsf{revc}([\alpha'](x), [\bar{a}](y)) && \text{(defn. } [\bar{a}]) \\ &= \mathsf{revc}(x, [\bar{\alpha'}]([\bar{a}](y))) && \text{(ind. hyp.)} \qquad \Box\end{aligned}$$

That the given interpretation orders R_P, is a consequence of the following lemma.

LEMMA 6.4. *Let* P *be a PCP instance that has no solution. For all* $\Phi, \Phi' \in \{f, g\}$, *and* $k \in \{0, 1\}$, *if*

$$h^k(\Phi(p, q, r, s, t)) \to \Phi'(p', q', r', s', t')$$

is a ground instance of a rule in R_P, *then*

$$[\Phi(p, q, r, s, t)] > [\Phi'(p', q', r', s', h(h(t)))]$$

holds.

Proof. By the form of the rules, and Lemma 6.3, we get

$$\mathsf{revc}([p],[q]) = \mathsf{revc}([p'],[q']) \tag{6.3}$$
$$\mathsf{revc}([r],[s]) = \mathsf{revc}([r'],[s']) \tag{6.4}$$

and so, by definition of bound,

$$\mathsf{bound}([p],[q],[r],[s]) = \mathsf{bound}([p'],[q'],[r'],[s'])\,. \tag{6.5}$$

For abbreviation let

$$\begin{aligned}
B &=_{\text{def}} (\mathsf{bound}([p],[q],[r],[s])+1)(\mathsf{revc}([p],[q])+\mathsf{revc}([r],[s])),\\
B' &=_{\text{def}} (\mathsf{bound}([p'],[q'],[r'],[s'])+1)(\mathsf{revc}([p'],[q'])+\mathsf{revc}([r'],[s'])),\\
lhs &=_{\text{def}} [\Phi(p,q,r,s,t)], \quad \text{and}\\
rhs &=_{\text{def}} [\Phi'(p',q',r',s',h(h(t)))]\,.
\end{aligned}$$

With that we get $b = b'$ by (6.3), (6.4), and (6.5), and so

$$\begin{aligned}
\lfloor \frac{lhs}{10} \rfloor &= [t]\cdot 156^{(\mathsf{len}_\Phi(\pi([p]),\pi([q]),\pi([r]),\pi([s]))+1+B)}\\
&= [t]\cdot 156^{(\mathsf{len}_\Phi(p,q,r,s)+1+B)} && \text{(Lemma 6.2)}\\
&\geq [t]\cdot 156^{(\mathsf{len}_{\Phi'}(p',q',r',s')+2+B)} && \text{(Prop. 6.2)}\\
&= [t]\cdot 156\cdot 156^{(\mathsf{len}_{\Phi'}(p',q',r',s')+1+B)}\\
&> ([t]\cdot 100+55)\cdot 156^{(\mathsf{len}_{\Phi'}(p',q',r',s')+1+B)} && ([t]\geq 1)\\
&= (([t]\circ 5)\circ 5)\cdot 156^{(\mathsf{len}_{\Phi'}(p',q',r',s')+1+B)} && (\text{defn. } \circ)\\
&= [h(h(t))]\cdot 156^{(\mathsf{len}_{\Phi'}(p',q',r',s')+1+B)} && (\text{defn. } [h])\\
&= [h(h(t))]\cdot 156^{(\mathsf{len}_{\Phi'}(p',q',r',s')+1+B')} && (B = B')\\
&= [h(h(t))]\cdot 156^{(\mathsf{len}_{\Phi'}(\pi([p']),\pi([q']),\pi([r']),\pi([s']))+1+B')} && \text{(Lemma 6.2)}\\
&= \lfloor \frac{rhs}{10} \rfloor
\end{aligned}$$

and so $lhs > rhs$. □

6.4. The interpretation is strictly monotonic

Every interpretation function $[0]$, $[1]$, $[\bar{0}]$, $[\bar{1}]$, $[h]$ is strictly monotonic in $\mathbb{N}_+$, by strict monotonicity of $\circ$.

Proposition 6.3. *$\circ$ and* revc *are strictly monotonic in each parameter.* bound *is (non-strictly) monotonic in each parameter.*

This leaves $[f]$ and $[g]$ for which strict monotonicity in each parameter is a "little more" difficult to prove. We need an auxiliary result:

Lemma 6.5. *If P is a PCP instance that has no solution then for all $x, y, z, w \in N_+$, ground terms t, and $\Phi \in \{f, g\}$*

$$\mathsf{len}_\Phi(\pi(x),\pi(y),\pi(z),\pi(w)) < \mathsf{bound}(x,y,z,w)\,.$$

PROOF. One easily establishes that

$$|\pi(x)| < \ell(x) \tag{6.6}$$

holds for all $x \in \mathbb{N}_+$, and that

$$\ell(\mathsf{revc}(x, y)) = 2\ell(x) + \ell(y) \tag{6.7}$$

holds for all $x, y \in \mathbb{N}_+$. Therefore:

$$\begin{aligned}
\mathsf{len}_\Phi(\pi(x), \pi(y), \pi(z), \pi(w)) &\leq 3 \cdot \min\{|\pi(x)| + |\pi(y)|, |\pi(z)| + |\pi(w)|\} && \text{(Prop. 6.1)} \\
&\leq 3 \cdot \min\{\ell(x) + \ell(y), \ell(z) + \ell(w)\} - 2 && \text{(6.6)} \\
&< 3 \cdot \min\{\ell(\mathsf{revc}(x, y)), \ell(\mathsf{revc}(z, w))\} - 2 && \text{(6.7)} \\
&= \mathsf{bound}(x, y, z, w) && \text{(defn. } \mathsf{bound}) \quad \square
\end{aligned}$$

Finally we arrive at:

LEMMA 6.6. *$[f]$ and $[g]$ are strictly monotonic in every parameter.*

PROOF. Obviously $[f]$ and $[g]$ are strictly monotonic in their last parameter. For the other parameters, let $x \leq x', y \leq y', z \leq z', w \leq w'$ where at least one of these inequalities is strict.

We observe that

$$\mathsf{revc}(x, y) + \mathsf{revc}(z, w) < \mathsf{revc}(x', y') + \mathsf{revc}(z', w') \tag{6.8}$$

holds by strict monotonicity of revc, and that

$$\mathsf{bound}(x, y, z, w) \leq \mathsf{bound}(x', y', z', w') \tag{6.9}$$

by monotonicity of bound. Let $\Phi \in \{f, g\}$, and let

$$\begin{aligned}
lhs &=_{\text{def}} [\Phi](x, y, z, w, u), \\
rhs &=_{\text{def}} [\Phi](x', y', z', w', u)\,.
\end{aligned}$$

Then we have:

$$\begin{aligned}
\log_{156}\lfloor \frac{lhs}{10u} \rfloor &= \mathsf{len}_\Phi(\pi(x), \pi(y), \pi(z), \pi(w)) + 1 \\
&\quad + (\mathsf{bound}(x, y, z, w) + 1)(\mathsf{revc}(x, y) + \mathsf{revc}(z, w)) \\
&< \mathsf{bound}(x, y, z, w) + 1 \\
&\quad + (\mathsf{bound}(x, y, z, w) + 1)(\mathsf{revc}(x, y) + \mathsf{revc}(z, w)) && \text{(Lemma 6.5)} \\
&= (\mathsf{bound}(x, y, z, w) + 1)(\mathsf{revc}(x, y) + \mathsf{revc}(z, w) + 1) \\
&\leq (\mathsf{bound}(x, y, z, w) + 1)(\mathsf{revc}(x', y') + \mathsf{revc}(z', w')) && \text{(6.8)} \\
&\leq (\mathsf{bound}(x', y', z', w') + 1)(\mathsf{revc}(x', y') + \mathsf{revc}(z', w')) && \text{(6.9)} \\
&\leq \mathsf{len}_\Phi(\pi(x'), \pi(y'), \pi(z'), \pi(w')) + 1 \\
&\quad + (\mathsf{bound}(x', y', z', w') + 1)(\mathsf{revc}(x', y') + \mathsf{revc}(z', w')) && (\mathsf{len}_\Phi(\ldots) \geq 0) \\
&= \log_{156}\lfloor \frac{rhs}{10u} \rfloor
\end{aligned}$$

So $\log_{156}\lfloor \frac{lhs}{10u} \rfloor < \log_{156}\lfloor \frac{rhs}{10u} \rfloor$, and so $lhs < rhs$. $\square$

7. Conclusions

A term rewriting system is called *totally terminating* if it is ordered by a strictly monotonic interpretation into a well-ordered set, and *ω-terminating*, if it is ordered by a strictly monotonic interpretation into the positive integer numbers. Both are special forms of termination of term rewriting systems, and the latter is a proper special case of the former. We have proven that the question whether a totally terminating term rewriting system is even ω-terminating is undecidable.

To this end we encoded instances, P, of the Post Correspondence Problem into term rewriting systems R_P in such a way that R_P is always totally terminating; and ω-terminating exactly if P has no solution.

Acknowledgements

Hans Zantema and Enno Ohlebusch checked the proofs, pointed out errors, and suggested several improvements.

References

Caron, A.-C. (1991). Linear bounded automata and rewrite systems: Influence of initial configurations on decision properties. In *Coll. Trees in Algebra and Programming*, LNCS 493, pages 74–89, Brighton. Springer.

Dauchet, M. (1992). Simulation of Turing machines by a regular rewrite rule. *Theoret. Comput. Sci.*, 103:409–420.

Dershowitz, N. (1987). Termination of rewriting. *J. Symbolic Computation*, 3(1&2):69–115. Corrigendum: 4, 409–410.

Dershowitz, N., Jouannaud, J.-P. (1991). Notations for rewriting. *Bull. EATCS*, 43:162–172.

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

Ferreira, M.C.F. (1995). *Termination of term rewriting—well-foundedness, totality, and transformations*. PhD thesis, University of Utrecht, NL.

Huet, G., Lankford, D.S. (1978). On the uniform halting problem for term rewriting systems. Technical Report 283, INRIA, Rocquencourt, FR, March.

Knuth, D.E., Bendix, P.B. (1970). Simple word problems in universal algebras. In Leech, J., ed., *Computational Problems in Abstract Algebra*, pages 263–297. Pergamon Press, Oxford.

Kurihara, M., Ohuchi, A. (1990). Modularity of simple termination of term rewriting systems. *J. IPS Japan*, 31(5):633–642.

Lescanne, P. (1994). On termination of one-rule rewrite systems. *Theoret. Comput. Sci.*, 132:395–401.

Manna, Z., Ness, S. (1970). On the termination of Markov algorithms. In *Proc. 3rd Intl. Conf. System Science*, pages 789–792, Honolulu, HI.

Meeussen, V., Zantema, H. (1993). Derivation lengths in term rewriting from interpretations in the naturals. In Wijshoff, H.A., ed., *Computing Science in the Netherlands*, pages 249–260, November. Also appeared as report RUU-CS-92-43, Utrecht University.

Plaisted, D. (1985). The undecidability of self-embedding for term rewriting systems. *Inform. Process. Lett.*, 20:61–64.

Post, E. (1946). A variant of a recursively unsolvable problem. *Bulletin of the American Mathematical Society*, 52.

Zantema, H. (1994). Termination of term rewriting: interpretation and type elimination. *J. Symbolic Computation*, 17(1):23–50.

Zantema, H. (1995). Total termination of term rewriting is undecidable. *J. Symbolic Computation*, 20:43–60.