On termination of one rule rewrite systems

Pierre LESCANNE
Centre de Recherche en Informatique de Nancy (CNRS)
and INRIA-Lorraine
Campus Scientifique, BP 239,
F54506 Vandœuvre-lès-Nancy, France

email: Pierre.Lescanne@loria.fr

Key-words: rewrite systems, well-foundedness

The indecidability of the termination of rewrite systems is usually proved by reduction to the halting of Turing machines. In particular, Dauchet proves the indecidability of the termination of one rule rewrite systems by coding Turing machines into one rule rewrite systems. Rewrite systems are a very simple model of computation and one may expect proofs in this model to be more straightforward and more satisfactory than those referring to the more complex model of Turing machines. In this paper we reduce the problem of termination of one rule rewrite systems to problems somewhat more related to rewrite systems namely to Post correspondence problems and to the termination of semi-Thue or monadic rewrite systems. Proofs we obtain this way are shorter and we expect other interesting applications from these codings.

Reduction of the termination of one rule rewrite systems to Post's correspondence problem

A correspondence system is a finite subset P of ordered pairs for some alphabet Σ, i.e., $P \equiv \{(\alpha_1, \beta_1\}, \ldots, (\alpha_n, \beta_n)\} \subseteq \Sigma^+ \times \Sigma^+$. A m-match of P is any string γ such $\gamma = \alpha_1 \alpha_{i_1} \ldots \alpha_{i_h} = \beta_1 \beta_{i_1} \ldots \beta_{i_h}$. Modified Post's correspondence problem is to determine, given a system, whether this system has a m-match. It is known [Pos47, LP81] that modified Post's correspondence system is undecidable even when Σ has two elements.

Consider a correspondence system $\{(\alpha_1, \beta_1\}, \ldots, (\alpha_n, \beta_n)\}$ of $\{0,1\}^+ \times \{0,1\}^+$ with $\|\alpha_1\| > \|\beta_1\|$, one can easily associate with it a system with $n+1$ rules which terminates if and only the Post's correspondence system has a m-match, namely,

$$\begin{array}{rcll} (\alpha_i x, \beta_i y, z) & \to & (x, y, z) & \qquad 1 \leq i \leq n \\ (\epsilon, \epsilon, \alpha_1 z) & \to & (\alpha_1 z, \alpha_1 z, \alpha_1 z) & \end{array}$$

Here we generalize this idea and since we are interested by a one rule rewrite system, we define a rewrite system

$$\begin{array}{l} A(C(0,1,\$),\ldots_{p \text{ times}}\ldots C(0,1,\$),(x_1 \cdot x_2 \cdot \ldots \cdot x_p \cdot x), \\ \quad C(0,1,\$),\ldots_{q \text{ times}}\ldots C(0,1,\$),(y_1 \cdot y_2 \cdot \ldots \cdot y_q \cdot y),\ a_1^1 \cdot \ldots \cdot a_{k_1}^1 \cdot z) \\ \qquad \to \\ B(t_1,\ldots,t_n,t) \end{array}$$

where $p = max\{\|\alpha_i\| \mid 1 \leq i \leq n\}, q = max\{\|\beta_i\| \mid 1 \leq i \leq n\}$. We also state $r = max(p,q)$. A, B and C are operators, A with arity $p+q+3$, B with arity $n+1$ and C with arity 3. nil, \$, 0 and 1 are four constants. Suppose $\alpha_i = a_1^i \ldots a_{k_i}^i$ with $k_i \leq p$ and $\beta_i = b_1^i \ldots b_{l_i}^i$ with $l_i \leq q$, then

$$t_i = A(u_1^i,\ldots,u_p^i,(x_{k+1} \cdot \ldots \cdot x_p \cdot x),v_1^i,\ldots,v_q^i,(y_{l+1} \cdot \ldots \cdot y_q \cdot y),a_1^1 \cdot \ldots \cdot a_{k_1}^1 \cdot z)$$

In that, if $1 \leq j \leq k$ and $a_j^i = 0$ then $u_j^i = C(x_j,1,\$)$, if $1 \leq j \leq k$ and $a_j^i = 1$ then $u_j^i = C(0,x_j,\$)$ and if $k < j \leq p$ then $u_j^i = C(0,1,\$)$. The same for the v_j^i's, namely if $1 \leq j \leq l$ and $b_j^i = 0$ then $v_j^i = C(y_j,1,\$)$, if $1 \leq j \leq l$ and $b_j^i = 1$ then $v_j^i = C(0,y_j,\$)$ and if $l < j \leq q$ then $v_j^i = C(0,1,\$)$. On the other hand,

$$t = A(w_1,\ldots,w_p,w'',w'_1,\ldots,w'_q,w'',w'')$$

where $w_i = C(0,1,x_i)$, $w'_i = C(0,1,y_i)$ and $w'' = a_1^1 \cdot \ldots \cdot a_{k_1}^1 \cdot z$. Call $LEFT$ the left-hand side of the rule and $RIGHT$ the right-hand side. A term s derived in the previous system from a ground instance of $LEFT$ has the following form: if at a position p there is an A, then above this position there are only B's and below this position there are only $C(a,b,c)$'s where a, b and c are in $\{0,1,\$\}$ or lists of 0, 1. The subterm $s_{|p}$ of s at position p is called a redex. A t_i-rewrite is a rewrite at one of the n first positions below a B and a t-rewrite is a rewrite at the last position below a B.

Example Consider the following correspondence problem [HU69]:

i	α_i	β_i
1	101	10
2	11	011
3	011	101

Figure 1 gives its associated one-rule rewrite system.

Lemma 1 *The following statements are equivalent*

(i) *The rewrite system* $LEFT \to RIGHT$ *does not terminate*

(ii) *There exist a substitution* σ *and an infinite derivation starting from* $LEFT\sigma$

(iii) *There exists* $s \xrightarrow{*} \Theta[s]$ *where* Θ *is a context,* $s = A(C(0,1,\$),\ldots,C(0,1,\$),\gamma,C(0,1,\$),\ldots,C(0,1,\$),\gamma,\gamma)$ *and* $\gamma = c_{i_1} \cdot \ldots \cdot c_{i_m} \cdot \$ \cdot \ldots_{r \text{ times}} \ldots \$ \cdot nil$.

(iv) $c_{i_1} \ldots c_{i_m}$ *is a solution of the Post's correspondence problem.*

$$
\begin{aligned}
&A(C(0,1,\$),C(0,1,\$),C(0,1,\$),x_1\cdot x_2\cdot x_3\cdot x,\\
&\quad C(0,1,\$),C(0,1,\$),C(0,1,\$),y_1\cdot y_2\cdot y_3\cdot x,\ 1\cdot 0\cdot 1\cdot z)\\
&\qquad\quad \rightarrow\\
&B(A(C(0,x_1,\$),C(x_2,1,\$),C(0,x_3,\$),x,\\
&\qquad C(0,y_1,\$),C(y_2,1,\$),C(0,1,\$),y_3\cdot y,\ 1\cdot 0\cdot 1\cdot z),\\
&\quad A(C(0,x_1,\$),C(0,x_2,\$),C(0,1,\$),x_3\cdot x,\\
&\qquad C(y_1,1,\$),C(0,y_2,\$),C(0,y_3,\$),y,\ 1\cdot 0\cdot 1\cdot z),\\
&\quad A(C(x_1,1,\$),C(0,x_2,\$),C(0,x_3,\$),x,\\
&\qquad C(0,y_1,\$),C(y_2,1,\$),C(0,y_3,\$),y,\ 1\cdot 0\cdot 1\cdot z)\\
&\quad A(C(0,1,\$),C(0,1,\$),C(0,1,\$),1\cdot 0\cdot 1\cdot z\\
&\qquad C(0,1,\$),C(0,1,\$),C(0,1,\$),1\cdot 0\cdot 1\cdot z,\ 1\cdot 0\cdot 1\cdot z))
\end{aligned}
$$

Figure 1: A one rule system associated with a Post's correspondence problem

Proof: The equivalence of the first two statements is a well-known result of termination. Suppose $z\sigma = \gamma$ and the derivation starting from $LEFT\sigma$ does not terminate then there are infinitely many t-rewrites are done. Indeed in each sequences of t_i-rewrites the size of the redexes decreases and t_i-rewrites alone cannot contribute to non termination. The redex associated with the first t-rewrite is

$$s = A(C(0,1,\$),\ldots,C(0,1,\$),\gamma,C(0,1,\$),\ldots,C(0,1,\$),\gamma,\gamma)$$

, it belongs to a term $\Phi[s]$ and if one considers only rewrites involved in creating the next t-redex (deep-only rewrites in Dauchet's terminology [Dau89]) it is easy to see that they build the term $\Phi[\Theta[s]]$, that $s \xrightarrow{*} \Theta[s]$ and that these rewrites contribute to build a m-match for the given Post's correspondence system, one occurrence of $c_{i_1}\ldots c_{i_m}$ rewrites to ϵ by the rules $\alpha_i \to \epsilon$ for $i_1 \ldots i_h$ and symmetrically the other rewrites also to ϵ using $\beta_i \to \epsilon$ with the same indices, when both reach ϵ, a m-match has been found. Reciprocally if there is a m-match $c_{i_1}\ldots c_{i_m}$ then the previous process creates an infinite sequence of rewrites starting from $LEFT(\$,\ldots,\$,\gamma,\$,\ldots,\$,\gamma,\gamma)$, where

$$\gamma = c_{i_1}\cdot\ldots\cdot c_{i_m}\cdot \$ \cdot \ldots_{r\text{ times}}\ldots \$ \cdot nil$$

□

From this lemma one gets Dauchet's theorem [Dau89].

Theorem 1 *Termination of orthogonal (i.e., left-linear and non-overlapping) one rule rewrite systems is undecidable.*

Reduction of the termination of one rule rewrite systems to the termination of monadic rewrite systems

We have seen in the previous section how to reduce the termination of a one rule rewrite system to the termination of a rewrite system where all the rewrites are done at the head. We use a similar method here to reduce termination of monadic rewrite systems to that of one rule rewrite systems. Thus suppose a monadic rewrite system $\mathsf{R} = \{\alpha_i x \to \beta_i x | 1 \leq i \leq n\}$ on $\{a_1, \ldots, a_m\}$ is given, it terminates exactly when the following system terminates.

$$\mathsf{R}' = \left\{ \begin{array}{rcll} (x, a_1^i \cdot \ldots \cdot a_{m_i}^i \cdot y, \mathsf{down}) & \to & (x, b_1^i \cdot \ldots \cdot b_{n_i}^i \cdot y, \mathsf{up}) & 1 \leq i \leq n \\ (x, a \cdot y, \mathsf{down}) & \to & (a \cdot x, y, \mathsf{down}) & \\ (a \cdot x, y, \mathsf{up}) & \to & (x, a \cdot y, \mathsf{up}) & \\ (\$, y, \mathsf{up}) & \to & (\$, y, \mathsf{down}) & \end{array} \right.$$

where $\alpha_i = a_1^i \ldots a_{m_i}^i$ and $\beta_i = b_1^i \ldots b_{n_i}^i$. The triple (x, y, d) should be understood as a position in a monadic term, x represents the part of the term above this position and y the part of the term below this position and d represents the direction in which the term is traversed. When going down one can either go further down or match a left-hand side, replace it by a right-hand side and go up. One goes up till the top is reached where one goes down. Suppose $p = max\{\|\alpha_i\| | 1 \leq i \leq n\}$, we associate with the monadic rewrite system R a one rule system:

$$\mathsf{R}'' = \{A(c, \ldots {\scriptstyle p \text{ times}} \ldots, c, \$, x' \cdot x, y_1 \ldots y_p \cdot y, \mathsf{down}, \mathsf{up}, d) \to B(t_1, \ldots, t_n, t'_1, t'_2, t'_3)\}$$

in the above rule $c = C(a_1, \ldots, a_m)$

$$t_i = A(c_1^i, \ldots, c_p^i, \$, x' \cdot x, b_1^i \ldots b_{n_i}^i \cdot y_{m_i+1} \ldots y_p y, d, \mathsf{up}, \mathsf{up})$$

where $c_j^i = C(a_1, \ldots, y_j, \ldots, a_n)$ for $1 \leq j \leq m_i$ and y_j lies at the k^{th} position if $a_j^i = a_k$ and $c_j^i = c$ for $m_i < j \leq p$

$$\begin{array}{rcl} t'_1 & = & A(c, \ldots, c, \$,\ y_1 \cdot x' \cdot x,\ y_2 \cdot \ldots \cdot y_p \cdot y,\ d, \mathsf{up}, \mathsf{down}) \\ t'_2 & = & A(c, \ldots, c, \$,\ x,\ x' \cdot y_1 \cdot \ldots \cdot y_p \cdot y,\ \mathsf{down}, d, \mathsf{up}) \\ t'_3 & = & A(c, \ldots, c, x',\ \$ \cdot nil,\ y_1 \cdot \ldots \cdot y_p \cdot y,\ \mathsf{down}, d, \mathsf{up}) \end{array}$$

Consider the same array as previously, but suppose now it describes a semi-Thue system. We then get the system of figure 2.

Lemma 2 R *terminates if and only if* R' *terminates if and only if* R'' *terminates*

Proof: Consider a term t. $t \to_R t'$ at position p if and only if

$$\begin{array}{l} (\$, r, \mathsf{down}) \to^*_{R'} (t_{p-1} \ldots t_1, t_p \ldots t_{k_p}, \mathsf{down}) \\ \qquad \to^*_{R'} (t_{p-1} \ldots t_1, t'_p \ldots t'_{k_p}, \mathsf{up}) \to^*_{R'} (\$, t', \mathsf{up}) \to^*_{R'} (\$, t', \mathsf{down}) \end{array}$$

t is irreducible if and only if $(\$, t, \mathsf{down})$ normalizes to $(\widetilde{t}, \$, \mathsf{down})$ where $\widetilde{t}$ is the mirror of t.

The arity of B is the same as the number of rules in R'. Each R''-rewrite (which takes place always under B) corresponds to an R'-rewrite (always at the top) hence the result. □

$$
\begin{array}{l}
A(C(0,1),C(0,1),C(0,1),\$,x'\cdot x,y_1\cdot y_2\cdot y_3\cdot y,\mathsf{down},\mathsf{up},d)\\
\qquad\to\\
B(A(C(0,y_1),C(y_2,1),C(0,y_1),\$,x'\cdot x,1\cdot 0\cdot y,d,\mathsf{up},\mathsf{up}),\\
\quad A(C(0,y_1),C(0,y_2),C(0,1),\$,x'\cdot x,0\cdot 1\cdot 1\cdot y_3\cdot y,d,\mathsf{up},\mathsf{up}),\\
\quad A(C(y_1,1),C(0,y_2),C(0,y_3),\$,x'\cdot x,1\cdot 0\cdot 1\cdot y,d,\mathsf{up},\mathsf{up}),\\
\quad A(C(0,1),C(0,1),C(0,1),\$,y_1\cdot x'\cdot x,y_2\cdot y_3\cdot y,d,\mathsf{up},\mathsf{down})\\
\quad A(C(0,1),C(0,1),C(0,1),\$,x,x'\cdot y_1\cdot y_2\cdot y_3\cdot y,\mathsf{down},d,\mathsf{up})\\
\quad A(C(0,1),C(0,1),C(0,1),x',\$\cdot nil,y_1\cdot y_2\cdot y_3\cdot y,\mathsf{down},d,\mathsf{up}))
\end{array}
$$

Figure 2: A one rule rewrite system simulating a monadic rewrite system

As a consequence of Huet and Lankford theorem [HL78] on indecidability of termination of monadic rewrite system one gets another proof of Theorem 1.

References

[Dau89] M. Dauchet. Simulation of Turing machines by a left-linear rewrite rule. In N. Dershowitz, editor, *Proceedings 3rd Conference on Rewriting Techniques and Applications, Chapel Hill (N.C., USA)*, volume 355 of *Lecture Notes in Computer Science*, pages 109–120. Springer-Verlag, April 1989.

[HL78] G. Huet and D. S. Lankford. On the uniform halting problem for term rewriting systems. Technical Report 283, Laboria, France, 1978.

[HU69] J. E. Hopcroft and J. D. Ullman. *Introduction to Automata Theory, Languages and Computation.* Addison-Wesley, Reading, Mass., USA, 1969.

[LP81] H. R. Lewis and Ch. H. Papadimitriou. *Elements of the theory of computation.* Prentice Hall Inc., 1981.

[Pos47] E. L. Post. Recursive unsolvability of a problem by Thue. *The Journal of Symbolic Logic*, 12(1):1–11, March 1947.