Deciding Confluence of Certain Term Rewriting Systems in Polynomial Time*

Ashish Tiwari
SRI International
333 Ravenswood Ave
Menlo Park, CA, U.S.A
tiwari@csl.sri.com

Abstract

We present a polynomial time algorithm for deciding confluence of ground term rewrite systems. We generalize the decision procedure to get a polynomial time algorithm, assuming that the maximum arity of a symbol in the signature is a constant, for deciding confluence of rewrite systems where each rule contains a shallow linear term on one side and a ground term on the other. The existence of a polynomial time algorithm for deciding confluence of ground rewrite systems was open for a long time and was independently solved only recently [4]. Our decision procedure is based on the concepts of abstract congruence closure [2] and abstract rewrite closure [12].

1. Introduction

The problem of checking confluence of an arbitrary term rewrite system is undecidable. This problem was shown to be decidable for the special cases of ground rewrite systems and left-linear right-ground rewrite systems in [5] (the ground case was also solved independently in [10]), but the resulting algorithms had an EXPTIME complexity. In this paper, we present (a) a polynomial time decision procedure to decide the confluence of ground term rewrite systems, and (b) a polynomial time decision procedure to decide the confluence of term rewrite systems containing rules with shallow linear terms on one side and ground terms on the other. We assume that the maximum arity of a function symbol is a constant in the second case.

The existence of a polynomial time decision procedure for these problems has been open for past several years[1]. Very recently a polynomial time decision procedure for the first problem was independently presented in [4].

Confluence and termination are two fundamental properties of rewrite systems, which respectively guarantee the existence of at most one and at least one normal form for each term. If interest is in reaching a normal form, confluence ensures that the order of rewrites does not matter. Although these properties are undecidable for general rewrite systems, these, and several others, are decidable for ground rewrite systems. Decidability of confluence for ground rewrite systems was shown using tree automata and ground tree transducers [10, 5]. These results can be generalized to left-linear right-ground rewrite systems [5] and to the full first-order theory of ground rewriting [6].

In our previous work [12], we introduced the concept of an abstract rewrite closure along the lines of an abstract congruence closure [2]. In essence, abstract rewrite closure represents certain kinds of ground tree transducers. Just as abstract congruence closure can be used to efficiently decide the congruence relation induced by a set of ground equations, an abstract rewrite closure efficiently decides the rewrite relation, or reachability, induced by a set of directed equations. In this paper, we use the concepts of an abstract congruence closure and abstract rewrite closure to obtain a polynomial time algorithm to decide the confluence of ground term rewrite systems. The abstract congruence closure and abstract rewrite closure concepts can be easily generalized to more general classes of rewrite systems, for example, to systems where shallow linear terms occur on one side and ground terms on the other. This observation leads to a polynomial time decision procedure to decide the confluence of these systems (assuming maximum arity of symbols in the signature is a constant) as well.

For example, let $\mathbb{R} = \{a \to f(a,b), f(a,b) \to f(b,a)\}$ be a ground rewrite system. We observe that the set $\mathbb{R}$ is not confluent as the two terms $f(b,a)$ and $f(f(b,a),b)$, though congruent modulo $\mathbb{R}$, are not both reducible to any common term via $\mathbb{R}$. In order to decide if $\mathbb{R}$ is confluent, our procedure first constructs an abstract rewrite closure for this set,

*This research was supported in part by the National Science Foundation under grants CCR-0082560 and CCR-0086096, and under NASA grant NAS1-00079.

[1] See Problem#12, submitted in 1991, in the RTA list of open problems at http://www.lri.fr/~rtaloop/index.html

followed by a congruence closure for the rewrite closure. Finally, we deduce that $\mathbb{R}$ is not confluent using a test on the two closures, see Example 1 for details.

1.1. Preliminaries

Let Σ be a set, called a *signature*, with an associated *arity function* $\alpha : \Sigma \rightarrow \mathbb{N}$. We define $\mathcal{T}(\Sigma)$ as the smallest set such that $f(t_1, \ldots, t_n) \in \mathcal{T}(\Sigma)$ whenever $f \in \Sigma$, $\alpha(f) = n$ and $t_1, \ldots, t_n \in \mathcal{T}(\Sigma)$. The elements of the sets Σ and $\mathcal{T}(\Sigma)$ are respectively called *function symbols* and *ground terms* (over Σ). Note that elements a in Σ for which $\alpha(a) = 0$, called *constants*, are included in the set $\mathcal{T}(\Sigma)$. The symbols $s, t, u, \ldots$, with possible subscripts, are used to denote terms; $f, g, \ldots$, function symbols; and $a, b, \ldots$, constants. The *size*, $\|t\|$, of a term $t = f(t_1, \ldots, t_n)$ is defined as $1 + \sum_{i=1}^{n} \|t_i\|$. We write $t[s]$ to indicate that a term t contains s as a subterm and (ambiguously) denote by $t[u]$ the result of replacing a particular occurrence of s in t by u.

An *(undirected) equation* is an unordered pair of terms, written $s \approx t$. A *directed equation* or *rule* is an ordered pair of terms, written $s \rightarrow t$. The size of an equation $s \approx t$ or a rule $s \rightarrow t$ is defined to be $\|s\| + \|t\|$. If E is a set of rules, then we define $E^- = \{s \rightarrow t : t \rightarrow s \in E\}$ and $E^{\pm} = E \cup E^-$. The *rewrite relation* $\rightarrow_E$ induced by a set of ground rules E is defined by: $u \rightarrow_E v$ if, and only if, u contains l as a subterm and v is obtained by replacing l by r in u, where $l \rightarrow r$ is in E. A set E of (ground) rules is called a *(ground) rewrite system*. The *size*, $\|E\|$, of a set E of ground equations and rules is the sum of the sizes of individual equations or rules in E. The cardinality of a set E is denoted by $|E|$.

If $\rightarrow$ is a binary relation, then $\leftarrow$ denotes its inverse, $\leftrightarrow$ its symmetric closure, $\rightarrow^+$ its transitive closure and $\rightarrow^*$ its reflexive-transitive closure. Thus, $\leftarrow_E$ and $\rightarrow_{E^-}$ denote identical relations. A set of rules E is *terminating* if there exists no infinite reduction sequence $s_0 \rightarrow_E s_1 \rightarrow_E s_2 \cdots$ of terms. A *proof* of $s \rightarrow_E^* t$ (in E) is a finite sequence $s = s_0 \rightarrow_E s_1,\ s_1 \rightarrow_E s_2,\ \cdots, s_{k-1} \rightarrow_E s_k = t (k \geq 0)$, which is usually written in abbreviated form as $s = s_0 \rightarrow_E s_1 \rightarrow_E \cdots \rightarrow_E s_k = t (k \geq 0)$.

Any irreflexive and transitive relation on the set $\mathcal{T}(\Sigma)$ of terms is called an *ordering*. An ordering is *well-founded*, or *Noetherian*, if there is no infinite sequence of terms $s_1, s_2, \cdots$ such that $s_1 \succ s_2$, $s_2 \succ s_3$, and so on. An ordering $\succ$ is *closed under contexts* if $u[s] \succ u[t]$ whenever $s \succ t$. A *reduction ordering* is a well-founded ordering which is also closed under contexts. If the rewrite relation induced by a ground rewrite system is contained in a reduction ordering, then the rewrite system is terminating.

A ground rewrite system is confluent if for all terms $s, t \in \mathcal{T}(\Sigma)$ whenever $s \leftrightarrow_E^* t$, there exists a term $u \in \mathcal{T}(\Sigma)$ such that $s \rightarrow_E^* u \leftarrow_E^* t$. If every rule $l \rightarrow r$ in E is such that l is not reducible by $E - \{l \rightarrow r\}$ and r is in E-normal form, then E is said to be *fully reduced*. A confluent and terminating rewrite system is called convergent. A fully reduced ground rewrite system is convergent [11].

2. Abstract Rewrite Closure

We shall assume that U is an infinite set of constants disjoint from Σ. We use c, d, with possible subscripts, to denote elements of U and K to denote finite subsets of U. An abstract rewrite closure [12] gives a succinct representation for the rewrite relation induced by a set of ground rewrite rules, much in the same way as an abstract congruence closure [2] does for the congruence relation induced by a set of ground equations.

Definition 1 *Let Σ be a signature and K be a set of constants disjoint from Σ. A* D-rule *(with respect to Σ and K) is a rewrite rule of the form $f(c_1, \ldots, c_k) \rightarrow c$ where $f \in \Sigma$ is a k-ary function symbol and $c_1, \ldots, c_k, c$ are constants in the set K. A rewrite rule of the form $c \rightarrow f(c_1, \ldots, c_k)$ will be called a* reverse D-rule. *A* C-rule *(with respect to K) is a rule $c \rightarrow d$, where c and d are constants in K.*

A set of D-rules and C-rules (with respect to Σ and K) is a specification of a bottom-up tree automaton transitions [3]. The set K represents "states" in the tree automaton and D-rules and C-rules represent regular and ϵ-transitions respectively. A constant c in K is said to *represent* a term $t \in \mathcal{T}(\Sigma \cup K)$ via the rewrite system E if $t \leftrightarrow_E^* c$.

A *ground tree transducer*, or GTT in short, is a pair of bottom-up tree automata defined over the same signature Σ and over the same set of states K, see [3]. It is well-known [5] that the rewrite relation induced by a set of ground rewrite system can be represented by a GTT. An abstract rewrite closure for a ground rewrite system is a GTT with certain restrictions. Intuitively, an abstract rewrite closure for a ground rewrite system $\mathbb{R}$ (over Σ) is a rewrite system RC over an extended signature $\Sigma \cup K$ such that RC conservatively extends the rewrite relation $\rightarrow_{\mathbb{R}}$ and every reachability proof in RC is representable as a normal-form proof, or valley proof, of some sort.

Definition 2 *Let Σ be a signature, K be a set of constants disjoint from Σ, and $\mathbb{R}$ be a set of ground rules (over $\mathcal{T}(\Sigma)$). A tuple (E, F, B) is said to be an* **(abstract) rewrite closure** *(with respect to Σ and K)* **for** *(the rewrite relation induced by) $\mathbb{R}$ if*

(i) E is a set of D-rules, F is a set of D-rules and C-rules, B is a set of reverse D-rules and C-rules such that each constant $c \in K$ represents some term $t \in \mathcal{T}(\Sigma)$ via E,

(ii) the rewrite systems $E \cup F$ and $E \cup B^-$ are terminating; and for all terms $s, t \in \mathcal{T}(\Sigma)$, if $s \rightarrow^*_{E^\pm \cup F \cup B} t$ then $s \rightarrow^*_{E \cup F} \circ \leftarrow^*_{E \cup B^-} t$, and

(iii) for all terms s and t in $\mathcal{T}(\Sigma)$, $s \rightarrow^*_{\mathbb{R}} t$ if and only if $s \rightarrow^*_{E^\pm \cup F \cup B} t$.

The tuple (E, F, B) defines a GTT [3]: the set $E \cup F$ defines the transitions of the first automaton and the set $E \cup B^-$ defines the transitions of the second automaton (over the same set K of "states"). Such a pair defines a binary relation $\rightarrow^*_{E \cup F} \circ \leftarrow^*_{E \cup B^-}$ on the set of ground terms.

For the sake of completeness, we reproduce here from [12], the set of inference rules (based on superposition and ordered chaining) to construct an abstract rewrite closure (E, F, B) for a ground rewrite system $\mathbb{R}$ over Σ. The inference rule system is more general and will be used to construct an abstract congruence closure in Section 3. The construction rules are parameterized by an infinite countable set U of new names and a total ordering $\succ_U$ on this set. The finite set $K \subset U$ is chosen by the procedure from the set U. The ordering $\succ_U$ restricted to K, denoted by $\succ_K$, is well-founded (on K).

We use $\succ$ to denote the recursive path ordering on $\mathcal{T}(\Sigma \cup K)$ generated using a precedence in which every $f \in \Sigma$ has higher precedence than any $c \in K$ and the precedence on K coincides with $\succ_K$. Note that a D-rule is oriented from left to right in this ordering. The ordering $\succ$ is a reduction ordering [7].

The transition rules operate on tuples (I, E, R), where I is a set of ground rules (over $\Sigma \cup K$), E is a set of D-rules (and C-rules, in case of abstract congruence closure construction) and R is a set of (reverse) D-rules and C-rules. The set R can be partitioned into the set $F = \{s \rightarrow t \in R : s \succ t\}$ of *forward* rules and the set $B = \{s \rightarrow t \in R : t \succ s\}$ of *backward* rules. In order to construct a rewrite closure for $\mathbb{R}$, we start in a state $(\mathbb{R}, \emptyset, \emptyset)$ and apply the following rules until the components of the state tuple stabilize. We will show that the final state is of the form $(\emptyset, E, R)$ and the tuple (E, F, B), where $R = F \cup B$, is an abstract rewrite closure for $\mathbb{R}$.

Extension flattens terms by introducing new constants, *Simplification* simplifies terms using rules in E, *Orientation* moves undirected equations and directed rules to E and R respectively, *Deletion* deletes trivial rules, and *Composition* simplifies right-hand sides of E-rules.

$$
\begin{array}{lll}
(I[s], E, R) & \vdash_{Ext} & (I[c], E \cup \{s \rightarrow c\}, R) \\
(I[u], E \cup \{u \rightarrow c\}, R) & \vdash_{Sim1} & (I[c], E \cup \{u \rightarrow c\}, R) \\
(I \cup \{u \approx c\}, E, R) & \vdash_{Ori} & (I, E \cup \{u \rightarrow c\}, R) \\
(I \cup \{u \rightarrow v\}, E, R) & \vdash_{Ori} & (I, E, R \cup \{u \rightarrow v\}) \\
(I \cup \{t \approx t\}, E, R) & \vdash_{Del} & (I, E, R) \\
(I \cup \{t \rightarrow t\}, E, R) & \vdash_{Del} & (I, E, R) \\
(I, E, R \cup \{t \rightarrow t\}) & \vdash_{Del} & (I, E, R) \\
(I, E \cup \{u \rightarrow c\}, R[u]) & \vdash_{Sim2} & (I, E \cup \{u \rightarrow c\}, R[c]) \\
(I, E \cup \{u \rightarrow c\}, R) & \vdash_{Com} & (I, E \cup \{u \rightarrow d\}, R)
\end{array}
$$

where $s \rightarrow c$ is a D-rule, $u \rightarrow c$ is either a D-rule or a C-rule with $u \succ c$, $u \rightarrow v$ is either a C-rule, D-rule, or a reverse D-rule, $t \in \mathcal{T}(\Sigma \cup K)$ is any term, $c \in U$ is a new constant in the *Ext*ension rule, and $c \rightarrow d \in E$ in the *Com*position rule. The two deduction rules, *Sup*erposition and *Cha*ining, are:

$$
\begin{array}{lll}
(I, E \cup \{s[t] \rightarrow d\}, R) & \vdash_{Sup} & (I, E \cup \{l \rightarrow r\}, R) \\
(I, E, R) & \vdash_{Cha} & (I, E, R \cup \{u \rightarrow v\})
\end{array}
$$

where $u \rightarrow v \in CP(R) \cup CP(E, R)$ and $l \rightarrow r$ is obtained by collapsing $s[t] \rightarrow d$ by $t \rightarrow c \in E$, i.e., either (i) $s[t] \neq t$, $l = s[c]$, and $r = d$; or, (ii) $s[t] = t$, $d \succ c$, $l = d$, and $r = c$. The set $CP(E, R)$ of critical pairs between rules in E and R is defined as: $CP(E, R) = \{f(\ldots, d, \ldots) \rightarrow c : f(\ldots, d', \ldots) \rightarrow c \in E, d \rightarrow d' \in B\} \cup \{c \rightarrow f(\ldots, d, \ldots) : f(\ldots, d', \ldots) \rightarrow c \in E, d' \rightarrow d \in F\}$. The set $CP(R)$ of critical pairs between rules in R is defined as: $CP(R) = \{t[d] \rightarrow c : d \rightarrow s \in B, t[s] \rightarrow c \in F\} \cup \{c \rightarrow t[d] : c \rightarrow t[s] \in B, s \rightarrow d \in F\}$. Here $F \cup B$ is a partition of R into forward and backward rules, as defined above.

Theorem 1 *Let $\mathbb{R}$ be a ground rewrite system over some signature Σ. Let m denote the maximum arity of any function symbol in Σ. Then, an abstract rewrite closure (E, F, B) (over $\Sigma \cup K$) for $\mathbb{R}$ can be constructed (using the inference rules given above with the set U and ordering $\succ_U$) in $O(N^3(m+1)^2)$ time, where $N = 2|\Sigma||K|^{m+1} + |K|^2$, such that*

1. the cardinality $|K|$ of the set $K \subset U$ is $O(\|\mathbb{R}\|)$,

2. the rewrite systems $E \cup F$ and $E \cup B^-$ are reducing with respect to $\succ$, and

3. the rewrite system E is fully reduced with $|E| = O(\|\mathbb{R}\|)$ and it contains no C-rules.

Proof. Starting in state $(\mathbb{R}, \emptyset, \emptyset)$, we apply the inference rules given above using the following strategy: (i) apply *Extension* and *Simplification1* exhaustively, applying the latter eagerly, (ii) apply *Deletion* and *Orientation* exhaustively, and finally (iii) apply *Chaining* and *Simplification2*, possibly followed by *Deletion*, until no new inferences can be done. This derivation ends in a final state $(\emptyset, E, R)$, and if $F \cup B$ is a partition of R into forward and backward rules, then the tuple (E, F, B) is an abstract rewrite closure for $\mathbb{R}$, see [12].

Now we count the number of inference steps applied: Step (i) involves application of at most $O(\|\mathbb{R}\|)$ inference rules as each rule eliminates at least one Σ-symbol from the input $\mathbb{R}$. This implies that at most $O(\|\mathbb{R}\|)$ equations are added to E and hence, $|K| = O(\|\mathbb{R}\|)$. Since *Simplification1* is eagerly applied, the set E is fully reduced and there are no *Superposition* steps applicable. The set E remains unchanged after Step (i) and hence property (3) holds. The

number of *Deletion* and *Orientation* step is bounded by $O(\|\mathbb{R}\|)$ since each such step removes one rule from the transformed set $\mathbb{R}$ of input rules. The number of steps of *Chaining* is bounded by the number of distinct D-rules, reverse D-rules, and C-rules, which is N. On any given $F \cup B$ rule, at most one *Simplification2* step can be applied and hence the number of *Simplification2* steps is $O(N)$. There are no *Composition* inferences possible.

We can check if an inference rule is applicable in $O(N^2(m+1))$ time (assuming trivial data structures to store the rules). An inference rule can be applied in $O(m+1)$ time. Thus, a maximal derivation can be constructed in $O(N^3(m+1)^2)$ time. ■

3. Abstract Congruence Closure

An *abstract congruence closure* [2] (with respect to Σ and K) is a set E_{CC} of D-rules and C-rules (w.r.t. Σ and K) such that E_{CC} is convergent and each constant $c \in K$ represents some term $t \in \mathcal{T}(\Sigma)$ via E_{CC}. An abstract congruence closure *for* a set $\mathbb{E}$ of ground equations (over $\Sigma \cup K$) is an abstract congruence closure E_{CC} (w.r.t. Σ and $K \cup K'$) such that for any two terms s and t (over $\Sigma \cup K$), $s \leftrightarrow^*_{\mathbb{E}} t$ iff $s \leftrightarrow^*_{E_{CC}} t$. In [2], we showed that an abstract congruence closure can be constructed in time polynomial in the size of the input $\mathbb{E}$. The construction of an abstract congruence closure requires a universe U of constants and an ordering $\succ_U$ on it as well.

Theorem 2 *Let (E, F, B) be an abstract rewrite closure (over $\Sigma \cup K$) for the ground rewrite system $\mathbb{R}$ with respect to the ordering $\succ_U$. A fully reduced abstract congruence closure E_{CC} (over $\Sigma \cup K$) for $E \cup F \cup B$ (with respect to the same ordering $\succ_U$) can be constructed in $O(N^2)$ time, where N is defined as in Theorem 1. The rewrite systems E_{CC} and $E \cup F \cup B^-$ are both reducing with respect to the ordering $\succ$ defined above.*

Proof. Starting from the state $(\emptyset, E \cup F \cup B^-, \emptyset)$ and applying the inference rules given above (see also [2]) exhaustively (using the ordering $\succ_K$), we reach a final state $(\emptyset, E_{CC}, \emptyset)$. It follows from the results in [2] that E_{CC} is a fully reduced abstract congruence closure for $E \cup F \cup B$ and this can be obtained in quadratic time. Since the first component is empty in the starting state, we cannot use *Extension* and hence no new constants are introduced in the derivation. Since we use the same ordering $\succ_U$, it is easy to see that the last claim is true as well. ■

4. Properties of Abstract Closures

For the rest of this paper, we fix the following notation: the set $\mathbb{R}$ denotes a ground rewrite system over a signature Σ, the tuple (E, F, B) is an abstract rewrite closure for $\mathbb{R}$ over the signature $\Sigma \cup K$ constructed using the ordering $\succ_U$ (as in Theorem 1), the set E_{CC} is a fully reduced abstract congruence closure for $E \cup F \cup B$ over the signature $\Sigma \cup K$ constructed using the same ordering (as in Theorem 2), the ordering $\succ$ is the rpo on $\mathcal{T}(\Sigma \cup K)$ obtained using the precedence $\succ_K$ as described above, and the constant $N = 2|\Sigma||K|^{m+1} + |K|^2$ is an upper bound on the total number of distinct D-rules, reverse D-rules, and C-rules over $\Sigma \cup K$.

Lemma 1 *For every $c \in K$, there exists a term $s \in \mathcal{T}(\Sigma)$ such that $s \rightarrow^*_E c$.*

Proof. Using Definition 2, we know that for every $c \in K$, there is a $s \in \mathcal{T}(\Sigma)$ such that $s \leftrightarrow^*_E c$. Since E is convergent and it contains no C-rules (by Property (3) in Theorem 1), the claim follows. ■

Lemma 2 *The number of rules in E_{CC} is $O(N)$ and for all terms $s, t \in \mathcal{T}(\Sigma \cup K)$, it is the case that $s \leftrightarrow^*_{E_{CC}} t$ if and only if $s \leftrightarrow^*_{E^{\pm} \cup F \cup B} t$.*

Proof. Construction of an abstract congruence closure guarantees that $|E_{CC}| \leq |E \cup F \cup B|$ and hence the first claim follows. The second claim follows from the definition of an abstract congruence closure [2]. ■

Lemma 3 *The rewrite system $\mathbb{R}$ is confluent (over $\mathcal{T}(\Sigma)$) if and only if the rewrite system $E^{\pm} \cup F \cup B$ is confluent (over $\mathcal{T}(\Sigma \cup K)$).*

Proof. Suppose $\mathbb{R}$ is confluent. Let $s, t \in \mathcal{T}(\Sigma \cup K)$ such that $s \leftrightarrow^*_{E^{\pm} \cup F \cup B} t$. Using Lemma 1, it follows that there exist terms s' and t' in $\mathcal{T}(\Sigma)$ such that $s' \rightarrow^*_E s$ and $t' \rightarrow^*_E t$. It follows that $s' \leftrightarrow^*_{E^{\pm} \cup F \cup B} t'$, and Property (iii) of Definition 2 yields $s' \leftrightarrow^*_{\mathbb{R}} t'$. Since $\mathbb{R}$ is confluent, we have $s' \rightarrow^*_{\mathbb{R}} \circ \leftarrow^*_{\mathbb{R}} t'$, which implies $s' \rightarrow^*_{E^{\pm} \cup F \cup B} \circ \leftarrow^*_{E^{\pm} \cup F \cup B} t'$ (using Property (iii) of Definition 2), and hence $s \rightarrow^*_{E^{\pm} \cup F \cup B} \circ \leftarrow^*_{E^{\pm} \cup F \cup B} t$.

Conversely, suppose that $E^{\pm} \cup F \cup B$ is confluent. If s and t are two terms in $\mathcal{T}(\Sigma)$ such that $s \leftarrow^*_{\mathbb{R}} \circ \rightarrow^*_{\mathbb{R}} t$, then $s \leftarrow^*_{E^{\pm} \cup F \cup B} \circ \rightarrow^*_{E^{\pm} \cup F \cup B} t$ and using confluence, we have $s \rightarrow^*_{E^{\pm} \cup F \cup B} u \leftarrow^*_{E^{\pm} \cup F \cup B} t$. If $u \notin \mathcal{T}(\Sigma)$, then we can lift u to a term u' in $\mathcal{T}(\Sigma)$ using Lemma 1 such that $s \rightarrow^*_{E^{\pm} \cup F \cup B} u' \leftarrow^*_{E^{\pm} \cup F \cup B} t$. Then, using Property (iii) of Definition 2, it follows that $s \rightarrow^*_{\mathbb{R}} \circ \leftarrow^*_{\mathbb{R}} t$. ■

Given an abstract congruence closure E_{CC}, we define a *signature* to be a term of the form $f(c_1, \ldots, c_n)$, where $f \in \Sigma$ and $c_1, \ldots, c_n$ are E_{CC}-irreducible constants in K. If $t = f(t_1, \ldots, t_n) \in \mathcal{T}(\Sigma \cup K)$ is a term, then we say $f(c_1, \ldots, c_n)$ is a *signature of* t (with respect to E_{CC}) if $f(c_1, \ldots, c_n)$ is a signature and $t_i \leftrightarrow^*_{E_{CC}} c_i$ for all i. Note that the symbol f could have arity 0, in which case $n = 0$ above. The total number of signatures is bounded

by $|\Sigma||K'|^m$, where m is the maximum arity of a function symbol in Σ and $K' \subseteq K$ is the set of all E_{CC}-irreducible constants. The following properties of signatures will be used in the subsequent proofs.

Proposition 1 *Let E_{CC} be an abstract congruence closure with respect to Σ and K. The following is true of signatures defined with respect to E_{CC}:*

1. A term $t \in \mathcal{T}(\Sigma)$ has at most one signature.

2. A term $t \in \mathcal{T}(\Sigma \cup K)$ such that $t \leftrightarrow^*_{E_{CC}} c$ has a signature.

3. If the two terms $s = f(s_1, \ldots, s_n)$ and $t = f(t_1, \ldots, t_n)$ (over $\Sigma \cup K$) have signatures, then these signatures are distinct iff $s_i \not\leftrightarrow^*_{E_{CC}} t_i$ for some i.

5. Confluence of Ground Rewrite Systems

Corresponding to each constant $c \in K$, we associate a set $IRRSIG(c)$ consisting of all signatures of $(E \cup F)$-irreducible terms, i.e.,

$$IRRSIG(c) = \{f(c_1, \ldots, c_n) : c \leftrightarrow^*_{E_{CC}} f(c_1, \ldots, c_n) \text{ and } f(c_1, \ldots, c_n) \text{ is a signature of an } (E \cup F)\text{-irreducible term } f(s_1, \ldots, s_n) \in \mathcal{T}(\Sigma \cup K)\}$$

and a set $IRRCON(c)$ consisting of all $(E \cup F)$-irreducible constants equivalent to c, i.e.,

$$IRRCON(c) = \{d \in K : d \leftrightarrow^*_{E_{CC}} c \text{ and } d \text{ is } (E \cup F)\text{-irreducible}\}.$$

Note that if c and d are two constants in the same congruence class (i.e., $c \leftrightarrow^*_{E_{CC}} d$), then $IRRSIG(c) = IRRSIG(d)$ and $IRRCON(c) = IRRCON(d)$. Hence, we can consider the sets $IRRSIG(c)$ and $IRRCON(c)$ as being defined on congruence classes, or equivalently, on E_{CC}-irreducible constants. In this case, the definitions above can be simplified by replacing the condition $f(c_1, \ldots, c_n) \leftrightarrow^*_{E_{CC}} c$ by $f(c_1, \ldots, c_n) \rightarrow_{E_{CC}} c$ and replacing $d \leftrightarrow^*_{E_{CC}} c$ by $d \rightarrow_{E_{CC}} c$. This follows from the facts that E_{CC} is fully-reduced and $f(c_1, \ldots, c_n)$ is a signature.

The main technical result characterizes confluence of $E^{\pm} \cup F \cup B$ in terms of the sets $IRRSIG(c)$ and $IRRCON(c)$.

Lemma 4 *Let Σ, K, $\mathbb{R}$, E, F, B, and E_{CC} be as fixed above. Then, $E^{\pm} \cup F \cup B$ is confluent iff the following three conditions are true for all $c \in K$ in E_{CC}-normal form:*

(a) the set $IRRSIG(c)$ contains at most one element, i.e., $|IRRSIG(c)| \leq 1$,

(b) if $IRRSIG(c) = \{f(c_1, \ldots, c_n)\}$, then for all $c' \in IRRCON(c)$, there is a rule $c' \rightarrow f(c'_1, \ldots, c'_n)$ in $E^- \cup B$ such that $c_i \leftrightarrow^*_{E_{CC}} c'_i$, and

(c) for all $d, e \in IRRCON(c)$, it is the case that $d \leftarrow^*_{E \cup B^-} \circ \rightarrow^*_{E \cup B^-} e$.

Proof. $\Rightarrow$: Suppose $E^{\pm} \cup F \cup B$ is confluent.

Proof of condition (a): Assume that $f(c_1, \ldots, c_m)$ and $g(d_1, \ldots, d_n)$ are two distinct signatures in $IRRSIG(c)$ that represent the $(E \cup F)$-irreducible terms $s = f(s_1, \ldots, s_m)$ and $t = g(t_1, \ldots, t_n)$ respectively. Then, $s \leftrightarrow^*_{E_{CC}} t$, and hence, $s \leftrightarrow^*_{E^{\pm} \cup F \cup B} t$. By confluence, $s \rightarrow^*_{E^{\pm} \cup F \cup B} \circ \leftarrow^*_{E^{\pm} \cup F \cup B} t$, and using Property (ii) in Definition 2 of rewrite closures

$$s \rightarrow^*_{E \cup F} \circ \leftarrow^*_{E \cup B^-} \circ \rightarrow^*_{E \cup B^-} \circ \leftarrow^*_{E \cup F} t.$$

Since s and t are $E \cup F$-irreducible, we have $s \leftarrow^*_{E \cup B^-} \circ \rightarrow^*_{E \cup B^-} t$, which implies $f = g$, $m = n$, and $s_i \leftrightarrow^*_{E_{CC}} t_i$ for all i. But, this is impossible since the signatures $f(c_1, \ldots, c_m)$ and $g(d_1, \ldots, d_n)$, of s and t respectively, are distinct (using Property (3) of Proposition 1).

Proof of condition (b): Let $s = f(s_1, \ldots, s_m)$ be the $(E \cup F)$-irreducible term represented by c. Since $s \leftrightarrow^*_{E_{CC}} c'$, it follows that $s \leftrightarrow^*_{E^{\pm} \cup F \cup B} c'$. Using the fact that $E^{\pm} \cup F \cup B$ is confluent, c' and s are $(E \cup F)$-irreducible, and (E, F, B) is a rewrite closure, we get

$$c' \leftarrow^*_{E \cup B^-} \circ \rightarrow^*_{E \cup B^-} f(s_1, \ldots, s_n).$$

Thus, the claim follows.

Proof of condition (c): Condition (c) follows directly from the definition of confluence, abstract rewrite closure, and noting that d and e are $(E \cup F)$-irreducible.

$\Leftarrow$: Suppose conditions (a)–(c) are true, but the set $E^{\pm} \cup F \cup B$ is not confluent. Let $\{s, t\}$ be a minimal witness (with respect to the multiset extension $\succ^m$ of the ordering $\succ$) to the non-confluence, i.e.,

$$s \leftrightarrow^*_{E_{CC}} t, \quad \text{but} \quad \neg\exists u : s \rightarrow^*_{E^{\pm} \cup F \cup B} u \leftarrow^*_{E^{\pm} \cup F \cup B} t.$$

The terms s and t are $(E \cup F)$-irreducible, for if they are not, then we can find a smaller witness. We distinguish the following cases.

1. $s = f(s_1, \ldots, s_n)$, $t = f(t_1, \ldots, t_n)$, and $s_i \leftrightarrow^*_{E_{CC}} t_i$ for all i: This cannot be the case since for some i, the pair $\{s_i, t_i\}$ will give a smaller witness.

2. $s = f(s_1, \ldots, s_m)$, $t = g(t_1, \ldots, t_n)$, and either $f \neq g$ or $s_i \not\leftrightarrow^*_{E_{CC}} t_i$ for some i: In this case there is a top rewrite step in the proof $s \leftrightarrow^*_{E_{CC}} t$. Therefore, we have

   $$s = f(s_1, \ldots, s_m) \leftrightarrow^*_{E_{CC}} c \leftrightarrow^*_{E_{CC}} g(t_1, \ldots, t_n) = t.$$

   Since the signatures of s and t are distinct (Property (2) and (3) of Proposition 1), it follows that $|IRRSIG(c)| \geq 2$, which contradicts condition (a).

3. Both s and t are constants in K: Condition (c) leads to a contradiction.

4. Exactly one of s and t is a constant in K: Without loss of generality, assume that $t = c'$ is a constant, $s = f(s_1, \ldots, s_m)$, and c is the E_{CC}-normal form of s (and c'). Clearly, $c' \in IRRCON(c)$ and the signature of s, say $f(c_1, \ldots, c_m)$, is in $IRRSIG(c)$. It follows from condition (b) that there exists a rule $c' \to f(c'_1, \ldots, c'_n) \in E^- \cup B$ such that $c'_i \leftrightarrow^*_{E_{CC}} c_i$. Now, note that $\{s, c\} \succ^m \{s_i, c'_i\}$, and therefore, $s_i \to^*_{E^{\pm} \cup F \cup B} u_i \leftarrow^*_{E^{\pm} \cup F \cup B} c'_i$. Putting all of these proofs together, we get a proof for

$$\begin{array}{lll} s & \to^*_{E^{\pm} \cup F \cup B} & f(u_1, \ldots, u_n) \\ & & f(c'_1, \ldots, c'_n) \end{array} \quad \begin{array}{l} \leftarrow^*_{E^{\pm} \cup F \cup B} \\ \leftarrow_{E^- \cup B} \end{array} \quad c'.$$

This leads to a contradiction.

This completes the proof of the lemma. ■

The final step of our proof consists of showing that the three conditions in Lemma 4 can be checked in polynomial time. We show that $IRRSIG(c)$ and $IRRCON(c)$ can be computed in polynomial time.

Lemma 5 *For each E_{CC}-irreducible constant $c \in K$, the sets $IRRSIG(c)$ and $IRRCON(c)$ can be computed in $O(N^3(m+1))$ time.*

Proof. Let c be in E_{CC}-normal form. Since E_{CC} is fully reduced, a constant $d \in K$ is in $IRRCON(c)$ if (1) $d \to_{E_{CC}} c$, and (2) d is $(E \cup F)$-irreducible. In time $O(|E \cup F||K|)$ we can determine, for all constants $c \in K$, if c is $(E \cup F)$-reducible. Hence, we can compute $IRRCON(c)$, for all c, in $O(|E_{CC}||K| + |E \cup F||K|)$ time.

A signature $f(c_1, \ldots, c_n)$ is in the set $IRRSIG(c)$ iff (1) $f(c_1, \ldots, c_n) \to c \in E_{CC}$ and either (2a) $f(c'_1, \ldots, c'_n)$ is $(E \cup F)$-irreducible for some $c'_1 \in IRRCON(c_1), c'_2 \in IRRCON(c_2), \ldots, c'_n \in IRRCON(c_n)$, or (2b) $IRRSIG(c_i)$ is non-empty for some i. We compute the sets $IRRSIG(c)$, for all E_{CC}-irreducible constants c, by initializing the sets $IRRSIG(c)$ by signatures that satisfy (1) and (2a) and finally adding signatures which satisfy (1) and (2b) (least fixed point computation). Note that there are at most $|K|$ iterations of the fixed point computation and each iteration takes $O(m|E_{CC}|)$ time. The initialization step takes $O(|E_{CC}||K|^m m|E \cup F|)$ time.

Thus, the total time taken is $O(|E_{CC}||K|^m m|E \cup F| + |E_{CC}||K|)$, which clearly is $O(N^3(m+1))$. ■

Using Lemma 5 it is easy to see that Conditions (a) and (b) in Lemma 4 can be checked efficiently.

Lemma 6 *Condition (c) of Lemma 4 can be checked in $O(N^2 + N^3 m)$ time.*

Proof. Let c be an E_{CC}-irreducible constant. Note that $|IRRCON(c)| \leq |K|$. Define the binary relation $\uparrow \subseteq K \times K$ by: $d \uparrow e$ iff $d \leftarrow^*_{E \cup B^-} \circ \to^*_{E \cup B^-} e$. We show how to compute $\uparrow$. Define $Src(d) = \{d' \in K : d' \to^*_{E \cup B^-} d\}$. Note that $d \in Src(d)$. Clearly, we can compute $Src(d)$, for all constants d, in time $O(|K|^2)$.

Now, $d \uparrow e$ is true if either (1) $Src(d) \cap Src(e)$ is nonempty, or (2) $d' \uparrow e'$ is true for some $d' \in Src(d)$ and $e' \in Src(e)$, or (3) there exists rules $f(d_1, \ldots, d_n) \to d$ and $f(e_1, \ldots, e_n) \to e$ in $E \cup B^-$ such that $d_i \uparrow e_i$ is true for all i.

Using the above characterization, the relation $\uparrow$ can be computed using a least fixed point computation again: we initialize the relation $\uparrow$ with pairs (d, e) that satisfy condition (1) and subsequently we add elements to this relation when either condition (2) or (3) holds. The initialization process takes $O(|K|^4)$ and the fixed point iterations take another $O(|K|^2(|K|^2 + |E \cup B|^2 m))$ time, which clearly is $O(N^2 + N^3 m)$. ■

Theorem 3 *If $\mathbb{R}$ is a ground term rewrite system, then the confluence of $\mathbb{R}$ can be decided in $O(N^3(m^2+1))$ time.*

Proof. We outline the complete algorithm here.

1. Construct an abstract rewrite closure (E, F, B) for $\mathbb{R}$ using a set $K \subset U$ of constants and an ordering $\succ_U$ over this set. This can be done in $O(N^3(m+1)^2)$ time (Theorem 1).

2. Construct an abstract congruence closure E_{CC} for $E \cup F \cup B$ using the same ordering $\succ_K$ in $O(N^2)$ time (Theorem 2).

3. Construct the sets $IRRSIG(c)$ and $IRRCON(c)$ for each constant $c \in K$ that is in E_{CC}-normal form. It follows from Lemma 5 that this step can be done in $O(N^3(m+1))$ time.

4. Check Conditions (a), (b), and (c) for each E_{CC}-irreducible constant c. If all conditions are satisfied, then $\mathbb{R}$ is confluent, otherwise it is not. It follows from Lemma 6 that this step can be done in $O(N^2 + N^3 m)$ time.

The correctness of the procedure follows from Lemma 4. This procedure runs in $O(N^3(m^2+1))$ time. ■

Corollary 1 *The confluence of a ground term rewrite system is decidable in polynomial time.*

Proof. Since a ground rewrite system can be transformed, while preserving confluence, into a rewrite system where the maximum arity of any function symbol is at most two [8, 4], we can assume that $m \leq 2$. Using Theorem 3, we get a $O(\|\mathbb{R}\|^9)$ time complexity procedure for deciding confluence of a GTRS $\mathbb{R}$. ■

We illustrate our decision procedure on two examples taken from [5].

Example 1 *Consider the rewrite system* $\mathbb{R} = \{a \to fab, fab \to fba\}$. *If* $U = \{c_1, c_2, \ldots\}$ *and* $\succ_U$ *is defined by* $c_1 \succ c_2 \succ \cdots$, *then the following illustrates the four steps used for deciding if* $\mathbb{R}$ *is confluent:*
1. An abstract rewrite closure for $\mathbb{R}$ *is:*

$$\begin{aligned} E &= \{a \to c_1,\ b \to c_2,\ fc_1c_2 \to c_3\} \\ F &= \{c_1 \to c_3\} \\ B &= \{c_3 \to fc_2c_1,\ c_3 \to fc_3c_2,\ c_3 \to fc_2c_3\} \end{aligned}$$

2. An abstract congruence closure for $E \cup F \cup B$ *is:*

$$\begin{aligned} E_{CC} = \{&a \to c_1,\ b \to c_2,\ fc_3c_2 \to c_3, \\ &c_1 \to c_3,\ fc_2c_3 \to c_3\} \end{aligned}$$

3. Constants in $K = \{c_1, c_2, c_3\}$ *that are* E_{CC}*-irreducible are* c_2 *and* c_3.

$$IRRSIG(c_2) = \emptyset, \quad IRRSIG(c_3) = \{fc_3c_2, fc_2c_3\}$$

4. Since $IRRSIG(c_3)$ *is not singleton, we conclude that the rewrite system* $\mathbb{R}$ *is not confluent.*

Example 2 *Consider the rewrite system* $\mathbb{R} = \{gfa \to fgfa, gfa \to ffa, ffa \to fa\}$. *Let* U *and* $\succ_U$ *be as in Example 1. The following illustrates the four steps used for deciding if* $\mathbb{R}$ *is confluent:*
1. An abstract rewrite closure for $\mathbb{R}$ *is:*

$$\begin{aligned} E &= \{a \to c_1,\ fc_1 \to c_2,\ gc_2 \to c_3\} \\ F &= \{fc_2 \to c_2,\ fc_3 \to c_2,\ gc_3 \to c_3\} \\ B &= \{c_3 \to fc_3,\ c_3 \to fc_2,\ c_3 \to c_2\} \end{aligned}$$

2. An abstract congruence closure for $E \cup F \cup B$ *is:*

$$\begin{aligned} E_{CC} = \{&a \to c_1,\ fc_1 \to c_3,\ gc_3 \to c_3, \\ &fc_3 \to c_3,\ c_2 \to c_3\} \end{aligned}$$

3. The E_{CC}*-irreducible constants in* $K = \{c_1, c_2, c_3\}$ *are* c_1 *and* c_3.

$$\begin{aligned} IRRCON(c_1) &= \{c_1\}, & IRRSIG(c_1) &= \emptyset \\ IRRCON(c_3) &= \{c_2, c_3\}, & IRRSIG(c_3) &= \emptyset \end{aligned}$$

4. Conditions (a) and (b) are vacuously true here. In order to check condition (c), we compute $Src(c_2)$ *and* $Src(c_3)$ *as* $Src(c_2) = \{c_2\}$ *and* $Src(c_3) = \{c_2, c_3\}$. *Since* $Src(c_2) \cap Src(c_3)$ *is nonempty, condition (c) is true as well. Hence, we conclude that* $\mathbb{R}$ *is confluent.*

6. Shallow-Linear Ground Rewrite Systems

In this section, we generalize the results of the previous sections to allow certain kinds of non-ground terms, specifically shallow and linear, on one of the two sides of rewrite rules in $\mathbb{R}$. We assume here that the input rewrite system $\mathbb{R}$ is such that a variable does not occur as left- or right-hand side of a rule in $\mathbb{R}$, since any shallow-linear ground rewrite system which violates this assumption is trivially confluent.

6.1. Definitions

Let $\mathcal{V}$ denote a denumerable set of *variables* disjoint from $\Sigma \cup U$. A term t in $\mathcal{T}(\Sigma \cup U, \mathcal{V})$ is said to be *linear* if any variable occurs at most once in t, and it is *shallow* if all variables occur at depth at most one. We say that a term rewrite system $\mathbb{R}$ is *shallow-linear ground* if for each rule $l \to r$ in $\mathbb{R}$, either l is a shallow and linear term and r is ground, or r is shallow and linear and l is ground. The size of a term, rule, and rewrite system is defined by suitably generalizing the definitions given previously for the ground case. The rewrite relation $\to_{\mathbb{R}}$ induced by a shallow-linear ground TRS is defined by: $u \to_{\mathbb{R}} v$ iff $u = u[l\sigma]$ contains $l\sigma$ as a subterm and $v = u[r\sigma]$ is obtained by replacing $l\sigma$ in u by $r\sigma$, for some rule $l \to r \in \mathbb{R}$ and substitution σ.

6.2. Abstract Rewrite Closure

Let Σ be a signature and K be a set of constants disjoint from Σ. A *D-rule* (with respect to Σ and K) is a rewrite rule of the form $f(\gamma_1, \ldots, \gamma_k) \to c$ where $f \in \Sigma$ is a k-ary function symbol, $c \in K$ is a constant, each $\gamma_i \in K \cup \mathcal{V}$ is either a constant in K or a variable in $\mathcal{V}$, and the term $f(\gamma_1, \ldots, \gamma_k)$ is linear. A rule $c \to f(\gamma_1, \ldots, \gamma_k)$ is a *reverse D-rule* if $f(\gamma_1, \ldots, \gamma_k) \to c$ is a D-rule. A rule $c \to d$, where c and d are constants in K is a *C-rule*.

Using these generalized definitions of D-rules, reverse D-rules, and C-rules, we can define an *abstract rewrite closure* for a shallow-linear ground TRS $\mathbb{R}$.

Definition 3 *A tuple* (E, F, B) *(over* $\mathcal{T}(\Sigma \cup K, \mathcal{V})$*) is an* **abstract rewrite closure for** *a shallow-linear ground rewrite system* $\mathbb{R}$ *(over* $\mathcal{T}(\Sigma, \mathcal{V})$*) if:*

(i) E *and* F *are sets of* D*-rules and* C*-rules,* B *is a set of reverse* D*-rules and* C*-rules (with respect to* Σ *and* K*), and each constant* $c \in K$ *represents some term* $t \in \mathcal{T}(\Sigma, \mathcal{V})$ *via* E,

(ii) the rewrite systems $E \cup F$ *and* $E \cup B^-$ *are terminating; and for all terms* $s, t \in \mathcal{T}(\Sigma, \mathcal{V})$, *if* $s \to^*_{E^\pm \cup F \cup B} t$ *then* $s \to^*_{E \cup F} \circ \leftarrow^*_{E \cup B^-} t$, *and*

(iii) for all terms s *and* t *in* $\mathcal{T}(\Sigma, \mathcal{V})$, $s \to^*_{\mathbb{R}} t$ *iff* $s \to^*_{E^\pm \cup F \cup B} t$.

We can construct an abstract rewrite closure for a shallow-linear ground TRS $\mathbb{R}$ using suitably generalized variants of the inference rules given before [13]. In particular, D-rules, C-rules, and reverse D-rules refer to the new definitions of these terms. The other differences are (i) the *Extension* rule is restricted to introducing only ground D-rules, (ii) the *Orientation* rule is used to move D-rules and reverse-D-rules containing variables from the first to the third component (for constructing a rewrite closure) and from first to the second (for constructing a congruence closure), (iii) the *Simplification* rules use standard *matching*

procedures to simplify a term, while (iv) the *Superposition* and *Chaining* rules use *unification* to compute critical pairs.

A crucial observation here is that the critical pairs generated by the *Superposition* and *Chaining* rules are always either D-rules, reverse D-rules, or C-rules. The linearity assumption ensures that variable chaining is not required for completeness and shallowness guarantees that terms can be flattened by *Extension*. The correctness argument for the inference rules is shown using proof simplification techniques [13]. If we consider two terms that are identical upto variable renaming (alpha conversion) as being equal, then the total number of possible D-rules, reverse D-rules, and C-rules, is bounded above by $N = 2|\Sigma|(|K|+1)^{m+1} + |K|^2$. The termination argument is identical to that of the ground case and the time complexity for constructing an abstract rewrite closure for a shallow-linear ground term rewrite system is $O(N^3(m+1)^2)$ following the arguments of Theorem 1, see [13]. We define the recursive path ordering $\succ$ with respect to which the rewrite system $E \cup F \cup B^-$ are reducing in the same way as before.

6.3. Abstract Congruence Closure

An abstract congruence closure can be defined and constructed using the generalized inference rules outlined above for abstract rewrite closure. For our purposes here, we only need to construct an abstract congruence closure for a set $E \cup F \cup B$ (over $\mathcal{T}(\Sigma \cup K, \mathcal{V})$) where (E, F, B) is a rewrite closure. As before, we can efficiently compute an abstract congruence closure E_{CC} for $E \cup F \cup B$ with the following properties [13]: (i) E_{CC} is a set of D-rules and C-rules (with respect to Σ and K) as defined above, (ii) every constant $c \in K$ represents some term $t \in \mathcal{T}(\Sigma, \mathcal{V})$ via E_{CC}, (iii) the equational theory $\leftrightarrow^*_{E \cup F \cup B}$ induced by $E \cup F \cup B$ is identical to the equational theory induced by E_{CC} over $\mathcal{T}(\Sigma \cup K, \mathcal{V})$, and (iv) the rewrite system E_{CC} is convergent and fully reduced.

The construction of an abstract congruence closure of a rewrite closure involves no *Extension* steps. Note that the critical pairs generated by *Superposition* are all D-equations or C-equations. The correctness of a fair derivation follows from standard results in term rewriting and termination argument is similar to that for computing the rewrite closure. The time complexity for computing a fully-reduced abstract congruence closure is, therefore, $O(N^3(m+1)^2)$, where N and m denote the same quantities as in the last subsection, see [13] for details.

6.4. Properties

If (E, F, B) is an abstract rewrite closure constructed by the strategy outlined in the proof of Theorem 1, then the set E would contain only ground D-rules and no C-rules.

Lemma 7 *Let E, F, B, E_{CC}, and $\mathbb{R}$ be as defined above. Then,*

1. for every $c \in K$, there exists a ground term $s \in \mathcal{T}(\Sigma)$ such that $s \rightarrow^*_E c$,

2. the number of rules in E_{CC} is $O(N)$, and

3. the rewrite system $\mathbb{R}$ is confluent (over $\mathcal{T}(\Sigma, \mathcal{V})$) if, and only if, the rewrite system $E^{\pm} \cup F \cup B$ is confluent (over $\mathcal{T}(\Sigma \cup K, \mathcal{V})$).

6.5. Confluence of shallow-linear ground TRS

A linear term of the form $f(\gamma_1, \ldots, \gamma_n)$, where $f \in \Sigma$, and each γ_i is either an E_{CC}-irreducible constant in K, or a variable, is called a *signature*. The *signature of a term* $f(s_1, \ldots, s_n)$ is a signature $f(\gamma_1, \ldots, \gamma_n)$ such that $\forall(i : \gamma_i \in K).\gamma_i \leftrightarrow^*_{E_{CC}} s_i$ and $\forall(i : \gamma_i \in \mathcal{V}).\gamma_i = s_i$. Note that Proposition 1 holds for the general case with these new definitions. The total number of signatures, upto variable renaming, is bounded by $|\Sigma||K'+1|^m$, where m is the maximum arity of a function symbol in Σ and $K' \subseteq K$ is the set of all E_{CC}-irreducible constants.

We define the sets $IRRSIG(c)$ and $IRRCON(c)$, where $c \in K$ is an E_{CC}-irreducible constant, as before but using this new definition of signature of a term.

A non-ground rewrite system $\mathbb{R}$ is confluent if for all terms $s, t \in \mathcal{T}(\Sigma, \mathcal{V})$, whenever $s \leftrightarrow^*_{\mathbb{R}} t$, there exists a term $u \in \mathcal{T}(\Sigma, \mathcal{V})$ such that $s \rightarrow^*_{\mathbb{R}} u \leftarrow^*_{\mathbb{R}} t$. For example, the shallow-linear ground rewrite system $\mathbb{R} = \{fa \rightarrow fx, fa \rightarrow a\}$ over $\Sigma = \{f, a\}$ is not confluent, although it is confluent over the term universe $\mathcal{T}(\Sigma)$.

Lemma 8 *Let (E, F, B) be an abstract rewrite closure (over signature $\Sigma \cup K$) for the shallow-linear ground rewrite system $\mathbb{R}$. Let E_{CC} be an abstract congruence closure for $E \cup F \cup B$ over the same signature. Then, $E^{\pm} \cup F \cup B$ is confluent iff the following three conditions are true for all $c \in K$ in E_{CC}-normal form:*

(a) the set $IRRSIG(c)$ contains at most one ground signature and no non-ground signatures,

(b) if $IRRSIG(c) = \{f(c_1, \ldots, c_n)\}$, then for all $c' \in IRRCON(c)$, there is a rule $c' \rightarrow f(c'_1, \ldots, c'_n)$ in $E^- \cup B$ such that $\forall i.\ c_i \leftrightarrow^*_{E_{CC}} c'_i$, and

(c) for all $d, e \in IRRCON(c)$, it is the case that $d \leftarrow^*_{E \cup B^-} \circ \rightarrow^*_{E \cup B^-} e$.

Proof. Suppose $E^{\pm} \cup F \cup B$ is confluent. If $f(\gamma_1, \ldots, \gamma_m)$ is a non-ground signature in $IRRSIG(c)$ of an $(E \cup F)$-irreducible term s, then s and $s\sigma$, where σ is any variable renaming such that $x \neq x\sigma$ for some x in s, are such that $s \leftrightarrow^*_{E_{CC}} s\sigma$ (since both are equivalent to c). Hence, $s \leftrightarrow^*_{E^{\pm} \cup F \cup B} s\sigma$, and by confluence, $s \rightarrow^*_{E^{\pm} \cup F \cup B} \circ \leftarrow^*_{E^{\pm} \cup F \cup B} s\sigma$. Following the argument from proof of

Lemma 4, we conclude that $s \leftarrow^*_{E \cup B^-} \circ \rightarrow^*_{E \cup B^-} s\sigma$, which implies $x = x\sigma$ for all variables x in s, which contradicts the assumption on σ. In case $IRRSIG(c)$ has two ground signatures, we can argue as in proof of Lemma 4 to get a contradiction. This proves condition (a). Conditions (b) and (c) are proved using the same argument as in the proof of Lemma 4.

Suppose conditions (a)–(c) are true, but the set $E^{\pm} \cup F \cup B$ is not confluent. Let $\{s, t\}$ be a minimal witness (with respect to the multiset extension $\succ^m$ of the ordering $\succ$) to the non-confluence, as in the proof of Lemma 4. The terms s and t are $(E \cup F)$-irreducible, and using a case split as before, we note that the argument for the first and third cases is as before. In the second case, if s and t are as in case (2) in proof of Lemma 4, then the signatures of s and t are distinct (Property (1) and (3) of Proposition 1) and therefore, either $|IRRSIG(c)| \geq 2$ or $IRRSIG(c)$ contains a non-ground signature, which contradicts condition (a). The argument for case (3) is identical. Finally, for Case 4, note that if $s, t = c'$ and c are as described in the other proof, then if the signature of s is ground, then we use the same argument. The case where signature of s is non-ground is impossible because of condition (a). ■

Finally, we show that the three conditions in Lemma 8 can be checked in polynomial time.

Lemma 9 *For each E_{CC}-irreducible constant $c \in K$, the sets $IRRSIG(c)$ and $IRRCON(c)$ can be computed in $O(N^3(m+1))$ time.*

Proof. The sets $IRRCON(c)$ can be constructed as in proof of Lemma 5. A signature $f(\gamma_1, \ldots, \gamma_n)$ (upto variable renaming) is in the set $IRRSIG(c)$ iff (1) $f(\gamma_1, \ldots, \gamma_n) \rightarrow c \in E_{CC}$, and (2) $f(\gamma_1, \ldots, \gamma_n)$ is not an instance of a left-hand side of a rule in $E \cup F$, and either (3a) $f(\gamma'_1, \ldots, \gamma'_n)$ is $E \cup F$-irreducible for some $\gamma'_1, \ldots, \gamma'_n$ such that $\forall (i : \gamma_i \in K).\ \gamma'_i \in IRRCON(c_i)$ and $\forall (i : \gamma_i \in \mathcal{V}).\ \gamma'_i = \gamma_i$, or (3b) for some i such that $\gamma_i \in K$, $IRRSIG(\gamma_i)$ is non-empty. We compute the sets $IRRSIG(c)$, for all E_{CC}-irreducible constants c, using the above definition as before. The time complexity of the procedure remains unchanged. ■

Lemma 10 *Condition (c) of Lemma 8 can be checked in $O(N^2 + N^3 m)$ time.*

Proof. The proof is a minor modification of the proof of Lemma 10. In particular, now $d \uparrow e$ is true if either (1) $Src(d) \cap Src(e)$ is nonempty, or (2) $d' \uparrow e'$ is true for some $d' \in Src(d)$ and $e' \in Src(e)$, or (3) there exists rules $f(\gamma_1, \ldots, \gamma_n) \rightarrow d$ and $f(\delta_1, \ldots, \delta_n) \rightarrow e$ in $E \cup B^-$ such that $\forall (i : \gamma_i \in K).\ \gamma_i \uparrow \delta_i$ is true and $\forall (i : \gamma_i \in \mathcal{V}).\ \delta_i \in \mathcal{V}$ is true. The complexity analysis remains unchanged. ■

We can now state the following result.

Theorem 4 *If $\mathbb{R}$ is a left-linear right-ground rewrite system, then the confluence of $\mathbb{R}$ can be decided in $O(N^3(m^2+1))$ time.*

Proof. Note that the only difference in the complexity analysis comes from the computation of the abstract congruence closure for the rewrite closure for $\mathbb{R}$, but this complexity is clearly bounded by that of construction of an abstract rewrite closure. ■

Example 3 *Consider the rewrite system $\mathbb{R} = \{fa \rightarrow ffa, fx \rightarrow fa\}$. If U and $\succ_U$ is as in Example 1, then the following illustrates the four steps used for deciding if $\mathbb{R}$ is confluent:*

1. An abstract rewrite closure for $\mathbb{R}$ is:

$$\begin{aligned} E &= \{a \rightarrow c_1,\ fc_1 \rightarrow c_2\} \\ F &= \{fx \rightarrow c_2\} \\ B &= \{c_2 \rightarrow fc_2\} \end{aligned}$$

2. An abstract congruence closure for $E \cup F \cup B$ is:

$$E_{CC} = \{a \rightarrow c_1,\ fx \rightarrow c_2\}$$

3. Constants in $K = \{c_1, c_2\}$ that are E_{CC}-irreducible are c_1 and c_2.

$$\begin{aligned} IRRCON(c_1) &= \{c_1\}, \quad IRRSIG(c_1) = \emptyset \\ IRRCON(c_2) &= \{c_2\}, \quad IRRSIG(c_2) = \emptyset \end{aligned}$$

4. Since the three conditions are easily verified, we conclude that the rewrite system $\mathbb{R}$ is confluent.

7. Related Work and Conclusion

We have shown that the problem of deciding confluence of ground term rewrite systems is in polynomial time. The algorithm obtained is quite simple and is based on the concepts of abstract congruence closure and abstract rewrite closure. We also showed that the same algorithm, with minor generalizations, also applies to the class of shallow-linear ground term rewrite systems.

Confluence was shown decidable in polynomial time for ground rewrite systems over signatures containing at most one unary function symbol and finitely many constants in [9]. For arbitrary signatures, the polynomial time decidability was independently demonstrated first in [4]. The approach used in that paper is based on transforming the input $\mathbb{R}$ using a curry transformation and a conservative introduction of new constants. These two steps also appear, respectively, in the proof of Corollary 1 and as the *Ext*ension rule in our paper. While the new definitions added by the *Ext*ension rule are treated as directed equations (part of E-component) in this paper, they are explicitly added bi-directionally (both $c \rightarrow t$ and $t \rightarrow c$ are added)

in [4]. An explicit closure of the rewrite relation over a term universe is computed next in [4]. We compute a closure under *ordered* chaining and superposition inferences in our approach. Consequently, whereas rewrite proofs can be made *increasing* in [4], our closure computation transforms rewrite proofs to so-called *valley* proofs. This difference is crucial, because it allows us to give a simple characterization of confluence (Lemma 4) using abstract rewrite and congruence closures, which can be proved by induction schema defined using the ordering $\succ$ (w.r.t. which the closures are constructed). On the contrary, the proof in [4] progresses by defining "stability" and "stabilizability" of top function symbol in a term and showing that these properties can be decided in polynomial time. Finally, [4] completes the proof by presenting some necessary conditions on confluence and assuming them, showing deep joinability of all left-hand sides of rules in $\mathbb{R}$ with themselves. The stability properties are related to the concept of "signatures" in our work and can also be decided using an abstract rewrite closure. The computation of deep joinability involves some fixed point computation similar in idea to the one used in proofs of Lemmas 5 and 6.

The first proofs of decidability of confluence for ground systems were based on tree-automata techniques and therefore, it is not surprising that abstract rewrite closure, which can be seen as a ground tree transducer, is central to our algorithm. The $O(N^3(m^2+1))$ time complexity computed in Theorem 3, and consequently the $O(\|\mathbb{R}\|^9)$ time complexity of Corollary 1, is based on simple arguments and need not be optimal even for the decision algorithm described in this paper. A more careful analysis of a particular implementation that uses appropriate data-structures and term indexing mechanisms can potentially improve the worst case time complexity. In case of shallow-linear ground rewrite systems, a curry transformation cannot be performed without changing the rewrite relation, and hence in this case, our procedure is in polynomial time only under the assumption that the maximum arity m is a constant.

The notion of an abstract congruence closure has been extended to handle signatures containing associative and commutative (AC) symbols [1]. Similarly, it is possible to extend the inference rules for computing rewrite closures to handle certain kinds of AC symbols in the signature [12]. Thus, we conjecture that the techniques in this paper can be used to obtain algorithms for deciding confluence of ground rewrite systems over such richer signatures. Additionally, our approach could also be used for other kinds non-ground term rewrite systems.

Acknowledgements. We thank Rakesh Verma for pointing us to [4], Guillem Godoy for pointing out an error in the second part of this paper, and the referees for their comments.

References

[1] L. Bachmair, I.V.Ramakrishnan, A. Tiwari, and L. Vigneron. Congruence closure modulo associativity and commutativity. In H. Kirchner and C. Ringeissen, editors, *Frontiers of Combining Systems, 3rd Intl Workshop FroCoS 2000*, pages 245–259, Nancy, France, Mar. 2000. Springer-Verlag. LNAI 1794.

[2] L. Bachmair, A. Tiwari, and L. Vigneron. Abstract congruence closure. *J. of Automated Reasoning*, 2002. To appear. Preliminary version appeared in CADE 2000, LNAI 1831.

[3] H. Comon, M. Dauchet, R. Gilleron, F. Jacquemard, D. Lugiez, S. Tison, and M. Tommasi. Tree automata techniques and applications. Available at `http://www.grappa.univ-lille3.fr/tata`, 1997.

[4] H. Comon, G. Godoy, and R. Nieuwenhuis. The confluence of ground term rewrite systems is decidable in polynomial time. In *Proc. 42nd Symp. Foundations of Computer Science FOCS*. IEE Comp. Soc. Press, 2001.

[5] M. Dauchet, T. Heuillard, P. Lescanne, and S. Tison. Decidability of the confluence of finite ground term rewrite systems. *Information and Computation*, 88:187–201, 1990. Also in Proc. IEEE Symposium on Logic in Computer Science 1987.

[6] M. Dauchet and S. Tison. The theory of ground rewrite systems is decidable. In *Proc of the 5th IEEE Symposium on Logic in Computer Science*, pages 242–248, Philadelphia, PA, June 1990. IEEE Computer Society Press.

[7] N. Dershowitz. Orderings for term-rewriting systems. *Theoretical Computer Science*, 17:279–301, 1982.

[8] P. J. Downey, R. Sethi, and R. E. Tarjan. Variations on the common subexpressions problem. *J. ACM*, 27(4):758–771, 1980.

[9] A. Hayrapetyan and R. Verma. On the complexity of confluence for ground rewrite systems. In *7th Biennial Bar-Ilan Intl. Symp. on the Found. of Artificial Intelligence, BISFAI*. Web-based Publication, 2001.

[10] M. Oyamaguchi. The Church-Rosser property for ground term rewriting systems is decidable. *Theoretical Computer Science*, 49:43–79, 1987.

[11] W. Snyder. A fast algorithm for generating reduced ground rewriting systems from a set of ground equations. *Journal of Symbolic Computation*, 15(7), 1993.

[12] A. Tiwari. Rewrite closure for ground and cancellative AC theories. In R. Hariharan and V. Vinay, editors, *Conference on Foundations of Software Technology and Theoretical Computer Science, FST&TCS '2001*, pages 334–346, Bangalore, India, 2001. Springer-Verlag. LNCS 2245.

[13] A. Tiwari. On the combination of equational and rewrite theories induced by certain term rewrite systems. Technical report, Computer Science Laboratory, SRI International, Menlo Park, CA 94025, 2002.