The algorithms for deciding some properties of finite convergent string rewriting systems (Theory and applications in computer algebra)

Author(s)
Mingchao, Guo; Lian, Li

Citation
数理解析研究所講究録 (1993), 848: 177-187

Issue Date
1993-09

URL
http://hdl.handle.net/2433/83652

Type
Departmental Bulletin Paper

Textversion
publisher

Kyoto University
The algorithms for deciding some properties of finite convergent string rewriting systems

Guo Mingchao, Li Lian
Lanzhou University, Gansu, China

Abstract. The following problems are decidable in $O(mn^2)$ time for convergent system $R$ on alphabet $A$, where $m = |A|$, $n$ is the length of $R$, i.e., the sum of all length of words appeared in $R$.

(1) Is the monoid presented by $R$ finite?
(2) How many elements in the monoid presented by $R$ if it is finite?

and a sofware is given to decide these properties and some other properties for such monoid.

1. String rewriting systems

Let $A$ be an alphabet, $R$ a subset of $A^* \times A^*$ which is called the set pf string rewriting rules. Elements in $R$ has the the form of $(x, y)$. Let $>$ be an order which satisfies that if $x > y$, then for arbitrary strings $w, z \in R$, $wxz > wyz$. A string rewriting system $RS$ is a double $(A, R)$ where $A$ is an alphabet and $R$ is a set of rewriting rules. An oriented $RS$ is triple $(A, R, >)$, for each rule $(x, y)$ in $R$, we have $x > y$ and denote as $x \rightarrow y$.

Some reduction relation on $A^*$ is defined as follows:

(a) $w_1 \rightarrow w_2$ iff there exists $z_1, z_2, (x, y) \in R$, such that $w_1 = z_1 x z_2$, $w_2 = z_1 y z_2$.

(b) $\rightarrow^*$ is the reflexive transitive closure of $\rightarrow$.

(c) $\leftrightarrow^*$ is the symmetric closure of $\rightarrow^*$.

A word $w$ is called irreducible if there is no word $z$ such that $w \rightarrow z$.

An ORS is Noetherian if there is no infinite chain $w_1 \rightarrow w_2 \rightarrow \cdots \rightarrow w_n \rightarrow \cdots$.

An ORS is confluent iff for arbitrary words $w_1, w_2, w_1 \leftrightarrow w_2$, there exists a word $z$ such that $w_1 \rightarrow^* z, w_2 \rightarrow^* z$.

An ORS is convergent iff it is Noetherian and confluent.

$G$ is a generating relation of a monoid on alphabet $A$. Let $\rho$ be the minimum congruence containing $G$, then

$$M \cong A^*/\rho.$$ 

$\rho = \leftrightarrow^*$ when $G$ is reviewed as a set of rules, so we call $A^*/\leftrightarrow^*$ is the monoid defined by ORS. When ORS is convergent, the following properties are decidable [2].
(1) Is the monoid presented by $R$ finite?

(2) How many elements are there the monoid presented by $R$ if it is finite?

(3) How is the multiplication table of the monoid presented by $R$ created?

(4) Is it a trivial monoid?

(5) Is it a group?

(6) Is it commutative?

(7) Is it a free monoid?

Whether an ORS is convergent is decidable [2]. If the ORS is not convergent, we can use the Knuth-Bendix convergent procedure to get an equivalent one in most case (But there indeed exists some ORS that has no convergent system). In the following discussion, we always suppose that ORS is convergent. About the properties of (4), (5), (6) and (7), F. Otto has given some algorithms in polynomial time [2, 4]. Now we turn our attention to the properties of (1), (2) and (3).

2. The decision of Properties (1), (2) and (3).

The number of equivalent classes is the number of elements in monoid when the ORS is convergent. And every class has exactly one irreducible element. The set of all irreducible words $IRR(R)$ is a regular language which can be accepted by a finite state automaton. The cardinal of $IRR(R)$ is equal to get the cardinal of monoid defined by the rewriting system. So we can count the cardinal to get the cardinal of the monoid. Some details are as follows.

Denote $\text{suffix}(x) := \{v|x = uv, u, v \in A^*\}$
$\text{dom}(R) := \{x|(x, y) \in R\}$
$\text{prefix}(R) := \{x|l = xy, y \in A^+, l \in \text{dom}(R)\}.$

We construct the automaton $FSA$ recognizing $IRR(R)$. The states of $FSA$ are those words which are proper prefixes of the left-side of the rules in $R$ where $e$ is the start state. All the states are final states. We denote the state set as $F$. The transitive function $\delta$ is constructed by following way.

$$\delta(w, a) = \begin{cases} 
\text{undefined} & \text{suffix}(wa) \cap \text{dom}(R) \neq \emptyset \\
s & s \in L = \{x|x \in \text{suffix}(wa) \cap \text{prefix}(R)\}, \\
& \text{and } y \in L - \{s\}, |s| > |y| 
\end{cases}$$

Lemma 1: $IRR(R) = L(FSA)$ [5].
Theorem 2: Let $FSA(Q, A, \delta, e, F)$ be the automaton on alphabet $A$ constructed as above. Then for the state $w \in F$ and $a \in A$, to determine the next state $\delta(w, a)$ needs $O(n)$ time, where $n = |R|$.

Proof: The key to determine the next state $\delta(w, a)$ is to find the suffix of $wa$ such that $wa$ is exactly a prefix of the left-side of a rule in $R$.

Suppose the word $u$ is the left-side of a rule, $S(u) = \{ \text{The longest word in prefix of } u \text{ and suffix of } wa \}$; obviously, $S(u) \neq \emptyset$.

If $u$ itself is a suffix of $wa$, then $\delta(w, a)$ is undefined, otherwise for arbitrary $u$ in the left-side of the rule, the problem to find the longest word $S(u)$ is a substring recognizing problem [1]. It takes $O(|u|)$ time. When $u$ runs all the left of the rule in $R$, it takes $O(n)$ time.

Theorem 3: $FSA$ as above can be constructed in $O(mn^2)$ time, where $m = |A|$, $n = |R|$.

Proof: The automaton has been constructed if we have the transitive function. For every state $w$ and $a \in A$, the next state $\delta(w, a)$ can be determined in $O(n)$ steps. The number of states is less than $n$, so for all states and some $a \in A$, to determine the next states needs $O(n^2)$ time. When a runs through $A$, it takes $O(mn^2)$.

Theorem 4: The monoid presented by $R$ is infinite iff the automaton $FSA$ recognizing $IRR(R)$ has a circle.

Proof: If the monoid is infinite, then there is an element $w \in IRR(R)$, $|w| > n$, where $n$ is the number of the states of $FSA$. For $FSA$ has only $n$ states, so $w$ must pass some state twice at least, i.e., $FSA$ contains a circle. Conversely, if $FSA$ has a circle, there exist states $q, r$ and a word $w = xyz$, which satisfies

$$\delta(e, x) = q, \quad \delta(q, y) = q, \quad \delta(q, z) = r.$$  

Then $FSA$ accepts all the words like $xy^kz$. The monoid presented by $R$ is infinite.

Algorithm 5:
INPUT: A finite automaton recognizing $IRR(R)$ on alphabet $A$.

begin $s_1 := \{ e \}$;
for $i := 1 \to n$ do
begin $S_2 := \emptyset$;
begin $S_2 := \bigcup_{p \in S_1, a \in A} \delta(p, a)$;
$S_1 := S_2$;
end;
If $S_2 = \emptyset$ then OUTPUT : The monoid is finite
else OUTPUT : The monoid is infinite
end.
Theorem 6: The algorithm above needs $O(mn)$ time to determine the existence of circle, where $m = |R|$ and $n$ is the number of states in automaton FSA.

Proof: clearly.

If the monoid presented by $R$ is finite, then FSA recognizing $IRR(R)$ has no circle, denote the number of words accepted by state $q$ as $N(q)$, obviously, $N(e) = 1$.

$R(q) = \{p \mid \exists a \in A : \delta(p, a) = q\}$.

$T(q, p) = |\{a \in A \mid \delta(p, a) = q\}|$.

Theorem 7: Let $FSA(Q, A, \delta, e, F)$ be a automaton recognizing $IRR(R)$ which has no circle. Then it satisfies

(a) For each $q$

$$N(q) = \begin{cases} 1 & q = e \\ \sum_{p \in R(q)} N(p)T(p, q) & q \neq e \end{cases}$$

(b) $|IRR(R)| = \sum_{q \in F} N(q)$

Proof: (a). denote $L(q)$ as the path length from start state to the state $q$. Now let us induce on $L(q)$.

$L(q) = 0$, then $q$ is start state. It is obviously that $N(q) = 1$.

If $L(q) < t$, the result is true. Let $L(q) = t$, every path from $e$ to $q$ must pass uniquely a state $p$ in $R(q)$, by the induction, the number of word from $e$ to $p$ is $N(p)$. So the number of words passing $p$ from $e$ to $q$ is $N(p)T(p, q)$, i.e., $N(q) = \sum_{p \in R(q)} N(p)T(p, q)$

(b) is clear.

Now we give the algorithm for calculating the order of the monoid.

Algorithm 8:

**INPUT**: Automaton $FSA(Q, A, \delta, e, F)$ accepting $IRR(R)$.

begin

$U := \{e\}$;

$N\{e\} := 1$;

$S := \{q \mid \exists p \in U, a \in A : \delta(p, a) = q\}$;

While $S \neq \emptyset$ do

\begin{verbatim}
begin $T := \emptyset$;
for each $q \in S$ do
if $R(q) \subseteq U$ then
begin $U := U \cup \{q\}$;
$N(q) := \sum_{p \in R(q)} N(p)T(p, q)$;
\end{verbatim}
\end{verbatim}

end
\[ T := T \cup \{ q \}; \]
\[ S := T; \]
\[ \text{end.} \]

**OUTPUT:** order := \( \sum_{q \in F} N(q) \).

Theorem 9: The algorithm 8 takes \( O(mn^2) \) time where \( m = |A|, n \) is the number of states of the automaton.

For constructing the multiplication table from a monoid, we can calculate the irreducible words of monoid at first. Denote \( W(p) := \{ x|\delta(e, x) := p \} \). The algorithm is similar with algorithm 8 and given as follows:

Algorithm 10:

**INPUT:** A automaton \( FSA(Q, A, \delta, e, F) \) recognizing \( IRR(R) \), where \( IRR(R) \) is finite.

begin
\[ W(e) := \{ 1 \}; \]
\[ U := \{ e \}; \]
\[ S := \{ q \mid \exists p \in U, a \in A : \delta(p, a) = q \}; \]
While \( S \neq \emptyset \) do
begin
\[ T := \emptyset; \]
for each \( p \in S \) do
if \( R(p) \subseteq U \) then
begin
\[ U := U \cup \{ p \}; \]
\[ W(p) := \bigcup_{p \in R(p), a \in A} \{ xa \mid x \in W(q), \delta(q, a) = p \}; \]
\[ T := T \cup \{ p \}; \]
end;
\[ S := T; \]
end.

**OUTPUT:** \( IRR(R) := \bigcup_{p \in F} W(p) \).

Theorem 11: If monoid presented by \( R \) is finite, we can calculating \( IRR(R) \) in polynomial \( O(mn^2t) \) time, where \( m = |a|, n \) is the number of states of \( FSA, t = |IRR(R)| \).

If the order of elements is \( t \) and the length of longest element is \( n \), to calculate the irreducible words for all production of pairwise elements needs \( O(n) \) time. For there is a linear time algorithm to get the irreducible word from a given word with length \( 2n \) [2]. So we can create the multiplication table in \( O(nt^2) \) time.
If the order of the monoid is too large, or it is infinite, to construct the whole multiplication table is impossible. But we can create a finite block of the table when the concrete elements is given. The following algorithm gives finite elements which lengths are less than $K$ in a monoid.

Algorithm 12:
INPUT: $FSA(Q, A, \delta, e, F)$ recognizing $IRR(R)$, an integer $k > 0$.

begin $S := \{e\}$;  
\( W := \{1\} \);
\( \text{While } (S \neq \emptyset) \) and \( (\text{loop} < k) \) do
begin $T := \emptyset$;
\( \text{loop} := \text{loop} + 1 \);
\( \text{For each } (x, q) \in S \) do
begin $M_{(x,q)} := \{(xa, p) | \delta(q, a) = p, a \in A\}$
\( \text{If } M_{(x,q)} \neq \emptyset \) then
begin $W := W \cup \{y | (y, p) \in M_{(x,q)}\}$;
$T := T \cup M_{(x,q)}$;
end;
$S := T$;
end;

OUTPUT: If $S = \emptyset$ then $IRR(R) := W$
\( \text{else } W \) is a irreducible words set whose element's length is less than $k$.

When $|R| = 1$, then algorithm 12 takes $O(k)$ time. When $|A| > 1$, then it takes $O(\frac{m^{k+1}-1}{m-1})$ time.

3. Some examples about monoid presented by rewriting systems

Based on the discussion above and [2, 4]. A software is designed and works well. It can get a convergent system from a generated relation and decide the properties of 1 to 7. The following are examples running on the software.
1. Generated relation:
\[
R = \{ aaab = ba \quad bbb = 1 \quad babb = aaaa \quad aaaba = b \quad baa = ab \quad aaaaaa = bababa \\
bbab = aa \quad ababa = bb \quad aabb = bba \quad aabab = baba \}
\]

Lexicographical order
Convergent system:
\[
R = \{ aabab = baba \quad aabb = bba \quad ababa = bb \quad bbab = aa \quad aaaaaa = babab \\
bba = ab \quad aaaba = b \quad babb = aaaa \quad bbb = 1 \quad aaaaab = ba \}
\]

Properties:

<table>
<thead>
<tr>
<th></th>
<th>trivial:</th>
<th>finite:</th>
<th>order:</th>
<th>commutative:</th>
<th>free monoid:</th>
<th>group:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NO</td>
<td>YES</td>
<td>21</td>
<td>NO</td>
<td>NO</td>
<td>YES</td>
</tr>
</tbody>
</table>

A finite multiplication table as following:

```
Semigroup Multiplication Table
```

<table>
<thead>
<tr>
<th></th>
<th>aaab</th>
<th>abba</th>
<th>abab</th>
<th>aaaa</th>
</tr>
</thead>
<tbody>
<tr>
<td>b</td>
<td>abab</td>
<td>aaaa</td>
<td>babab</td>
<td>aaba</td>
</tr>
<tr>
<td>ba</td>
<td>bba</td>
<td>aa</td>
<td>aaa</td>
<td>aaab</td>
</tr>
<tr>
<td>bb</td>
<td>babab</td>
<td>aaba</td>
<td>aaab</td>
<td>abba</td>
</tr>
<tr>
<td>aa</td>
<td>abab</td>
<td>bab</td>
<td>babab</td>
<td>1</td>
</tr>
<tr>
<td>ab</td>
<td>baba</td>
<td>babab</td>
<td>1</td>
<td>b</td>
</tr>
<tr>
<td>aba</td>
<td>abba</td>
<td>aaa</td>
<td>aaaa</td>
<td>ba</td>
</tr>
<tr>
<td>abb</td>
<td>1</td>
<td>b</td>
<td>ba</td>
<td>bab</td>
</tr>
<tr>
<td>aaa</td>
<td>aaba</td>
<td>baba</td>
<td>abb</td>
<td>a</td>
</tr>
<tr>
<td>aab</td>
<td>bb</td>
<td>1</td>
<td>a</td>
<td>ab</td>
</tr>
<tr>
<td>bba</td>
<td>a</td>
<td>ab</td>
<td>aba</td>
<td>abab</td>
</tr>
<tr>
<td>bab</td>
<td>aaa</td>
<td>aaab</td>
<td>b</td>
<td>bb</td>
</tr>
<tr>
<td>babab</td>
<td>aaaa</td>
<td>aba</td>
<td>aab</td>
<td>bba</td>
</tr>
<tr>
<td>aaba</td>
<td>bab</td>
<td>aaaa</td>
<td>aaaa</td>
<td>aba</td>
</tr>
<tr>
<td>aaaa</td>
<td>b</td>
<td>bb</td>
<td>bba</td>
<td>aa</td>
</tr>
<tr>
<td>aaab</td>
<td>abb</td>
<td>a</td>
<td>aa</td>
<td>aab</td>
</tr>
<tr>
<td>abba</td>
<td>aa</td>
<td>aab</td>
<td>aaba</td>
<td>baba</td>
</tr>
<tr>
<td>abab</td>
<td>aaaa</td>
<td>ba</td>
<td>ab</td>
<td>abb</td>
</tr>
<tr>
<td>aaaa</td>
<td>ab</td>
<td>aab</td>
<td>abba</td>
<td>aaa</td>
</tr>
<tr>
<td>babab</td>
<td>aab</td>
<td>bba</td>
<td>bab</td>
<td>aaaa</td>
</tr>
</tbody>
</table>
```
2.
Generated relation:
\[ R = \{ \text{abb} = \text{aa} \quad \text{ababababab} = \text{b} \quad \text{ab} = \text{ba} \} \]

Lexicographical order

Convergent system:
\[ R = \{ \text{ba} = \text{ab} \quad \text{aaaaaaa} = \text{b} \quad \text{abb} = \text{aa} \quad \text{bbb} = \text{ab} \} \]

Properties:

trivial: NO
finite: YES
order: 19
commutative: YES
free monoid: NO
group: NO

A finite multiplication table as following:

```
Semigroup Multiplication Table

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>b</th>
<th>a</th>
<th>ab</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>b</td>
<td>a</td>
<td>ab</td>
</tr>
<tr>
<td>b</td>
<td>b</td>
<td>bb</td>
<td>ab</td>
<td>aa</td>
</tr>
<tr>
<td>a</td>
<td>a</td>
<td>ab</td>
<td>aa</td>
<td>aab</td>
</tr>
<tr>
<td>ab</td>
<td>ab</td>
<td>aa</td>
<td>.aab</td>
<td>aaa</td>
</tr>
<tr>
<td>aa</td>
<td>aa</td>
<td>aab</td>
<td>aaa</td>
<td>aab</td>
</tr>
<tr>
<td>bb</td>
<td>bb</td>
<td>ab</td>
<td>aa</td>
<td>aab</td>
</tr>
<tr>
<td>aab</td>
<td>aab</td>
<td>aaa</td>
<td>aaba</td>
<td>aaaa</td>
</tr>
<tr>
<td>aaaa</td>
<td>aaaa</td>
<td>aaaa</td>
<td>aaaaab</td>
<td>aaaa</td>
</tr>
<tr>
<td>aaaaaab</td>
<td>aaaaaa</td>
<td>aaaaab</td>
<td>aaaaab</td>
<td>aaaaab</td>
</tr>
<tr>
<td>aaaaaa</td>
<td>aaaaaa</td>
<td>aaaaab</td>
<td>aaaaab</td>
<td>aaaaab</td>
</tr>
<tr>
<td>aaaaaaa</td>
<td>aaaaaaa</td>
<td>aaaaab</td>
<td>aaaaab</td>
<td>aaaaab</td>
</tr>
<tr>
<td>aaaaaaaa</td>
<td>aaaaaaaa</td>
<td>aaaaab</td>
<td>aaaaab</td>
<td>aaaaab</td>
</tr>
<tr>
<td>aaaaaaaab</td>
<td>aaaaaaab</td>
<td>aaaaaa</td>
<td>aaaaaa</td>
<td>aaaaab</td>
</tr>
<tr>
<td>aaaaaaaab</td>
<td>aaaaaaab</td>
<td>aaaaaa</td>
<td>aaaaaa</td>
<td>b</td>
</tr>
<tr>
<td>aaaaaaaa</td>
<td>aaaaaaaa</td>
<td>aaaaaaab</td>
<td>b</td>
<td>bb</td>
</tr>
<tr>
<td>aaaaaaaab</td>
<td>aaaaaaab</td>
<td>b</td>
<td>bb</td>
<td>ab</td>
</tr>
</tbody>
</table>
```
3.
Generated relation:

\[ R = \{aba = abab \quad cbab = bcbacb \quad cbcb = bc \quad ca = ac \quad aa = 1 \quad bb = 1 \quad cc = 1 \} \]

Lexicographical order

Convergent system:

\[ R \text{ is convergent.} \]

Properties:

trivial: \quad NO

finite: \quad NO

commutative: \quad NO

free monoid: \quad NO

group: \quad YES

A finite multiplication table as following:

\[ \begin{array}{cccc}
\text{Semigroup Multiplication Table} \\
\hline
\text{bab} & \text{bc} & \text{abc} & \text{aba} \\
\hline
\text{acbabc} & \text{ac} & \text{acbabc} & \text{cbac} & \text{cb} \\
\text{acbabc} & \text{acbabc} & \text{acbabc} & \text{cbabcb} & \text{cbabca} \\
\text{acbabc} & \text{acbabc} & \text{acbabc} & \text{cbabcb} & \text{cbabca} \\
\text{abcabc} & \text{abc} & \text{abcabc} & \text{ab} & \text{abc} \\
\text{abcabc} & \text{abc} & \text{abcabc} & \text{ab} & \text{abc} \\
\text{abcabc} & \text{abc} & \text{abcabc} & \text{ab} & \text{abc} \\
\text{abcabc} & \text{abc} & \text{abcabc} & \text{ab} & \text{abc} \\
\text{bababc} & \text{bacb} & \text{bababc} & \text{bab} & \text{abab} \\
\end{array} \]
4.

Generated relation:
\[ R = \{ cc = 1 \quad bb = 1 \quad aa = 1 \quad ca = ac \quad bab = aba \quad cbacbacb = bcbacbcb \quad cbcb = bcbcb \quad cbcbac = cbcbca \quad cbacbacba = cbacbacba \} \]

Lexicographical order

Convergent system:

\( R \) is convergent.

Properties:

- trivial: NO
- finite: YES
- order: 120
- commutative: NO
- free monoid: NO
- group: YES

A finite multiplication table as following:

<table>
<thead>
<tr>
<th></th>
<th>cb</th>
<th>cbc</th>
<th>cba</th>
<th>bac</th>
</tr>
</thead>
<tbody>
<tr>
<td>bacbca</td>
<td>abacbba</td>
<td>abacbabc</td>
<td>abacbcb</td>
<td>babacbca</td>
</tr>
<tr>
<td>bacbcb</td>
<td>bacbacbc</td>
<td>bcabcacb</td>
<td>bcabcbca</td>
<td>bacbcb</td>
</tr>
<tr>
<td>bcacbc</td>
<td>bcabc</td>
<td>bcb</td>
<td>bc</td>
<td>bcabcbca</td>
</tr>
<tr>
<td>cbabcba</td>
<td>cbabcab</td>
<td>cbabcacb</td>
<td>cbabcacbca</td>
<td>cbabcba</td>
</tr>
<tr>
<td>abacbc</td>
<td>abcba</td>
<td>abcba</td>
<td>abcb</td>
<td>abacbac</td>
</tr>
<tr>
<td>abacbc</td>
<td>bcacb</td>
<td>bacb</td>
<td>bacba</td>
<td>abacbac</td>
</tr>
<tr>
<td>acbacba</td>
<td>abcba</td>
<td>abcba</td>
<td>abc</td>
<td>acbacbca</td>
</tr>
<tr>
<td>acbacbc</td>
<td>cbacb</td>
<td>cba</td>
<td>cbabc</td>
<td>acbacba</td>
</tr>
<tr>
<td>abcabcba</td>
<td>babcab</td>
<td>babcabac</td>
<td>babcabac</td>
<td>abcabc</td>
</tr>
<tr>
<td>ababcabc</td>
<td>ababcabc</td>
<td>ababcabc</td>
<td>ababcabc</td>
<td>ababc</td>
</tr>
<tr>
<td>abcabcba</td>
<td>abcabcba</td>
<td>abcabcba</td>
<td>abcabcba</td>
<td>abcabcba</td>
</tr>
<tr>
<td>acbacbc</td>
<td>abcabc</td>
<td>abcabc</td>
<td>abcabc</td>
<td>abcabc</td>
</tr>
<tr>
<td>cbacbc</td>
<td>bcabc</td>
<td>bcabc</td>
<td>bcabc</td>
<td>bcabc</td>
</tr>
<tr>
<td>bcacbc</td>
<td>babcabac</td>
<td>babcabac</td>
<td>babcabac</td>
<td>babcabac</td>
</tr>
<tr>
<td>bacacbc</td>
<td>babcab</td>
<td>babcab</td>
<td>babcab</td>
<td>babcab</td>
</tr>
</tbody>
</table>
References


