ELSEVIER
Theoretical Computer Science 289 (2002) 853–859
Theoretical Computer Science
www.elsevier.com/locate/tcs

Note

Algorithms for determining the smallest number of nonterminals (states) sufficient for generating (accepting) a regular language R with $R_1 \subseteq R \subseteq R_2$ for given regular languages R_1, R_2

Kosaburo Hashiguchi

Department of Information Technology, Faculty of Engineering, Okayama University, 1-1 Tsushima-kana, 3-chome, Okayama 700-0082, Japan

Received October 2000; revised July 2001; accepted August 2001
Communicated by M. Crochemore

Abstract

Given two regular languages R_1 and R_2 with $R_1 \subseteq R_2$, one can effectively determine the number of nonterminals in a nonterminal-minimal (generalized) right linear grammar generating a regular language R with $R_1 \subseteq R \subseteq R_2$, and the number of states in a state-minimal (generalized) nondeterministic finite automaton accepting a regular language R with $R_1 \subseteq R \subseteq R_2$.

Keywords: Nonterminal-minimal; Regular language; Right linear grammar; Algorithm

1. Introduction

A (generalized) right linear grammar (in short, an rl grammar) G is a quadruple, $\langle V, \Sigma, P, S \rangle$, where V is a finite set of nonterminals, Σ is a finite alphabet, P is a finite set of production rules, and $S \in V$ is the start symbol. Here, any rule in P is of one of the following forms: $A \to uB$ or $A \to v$ for some $A, B \in V$ and $u, v \in \Sigma^*$. $L(G)$ denotes the language generated by G, and $n(G)$ denotes the cardinality of V, $|V|$. For any regular language R, we define $n(R)$ by $n(R) = \min\{n(G) \mid G$ is an rl grammar generating $R\}$. An rl grammar G is R-nonterminal-minimal if $L(G) = R$ and $n(G) = n(R)$.

Ibarra [3] posed the following problem.

Problem A. Is it decidable to determine $n(R)$ for any given regular language R?

E-mail address: hasiguti@kiso.it.okayama-u.ac.jp (K. Hashiguchi).

PII: S0304-3975(01)00324-3

In [2], the author presents an affirmative answer to this problem by presenting an algorithm for determining $n(R)$ for any given regular language R.

In this note, we shall study the following problem. Let R_1 and R_2 be two regular languages. We define $n(R_1,R_2)$ by $n(R_1,R_2)=\min\{n(R)\mid R$ is a regular language with $R_1\subseteq R\subseteq R_2\}$ if $R_1\subseteq R_2$ and $n(R_1,R_2)=\infty$ otherwise.

An rl grammar G is (R_1,R_2)-nonterminal-minimal if $R_1\subseteq L(G)\subseteq R_2$ and $n(G)=n(R_1,R_2)$.

Problem B. Is it decidable to determine $n(R_1,R_2)$ for any given two regular languages R_1,R_2?

The main result of this note is Theorem 2.1 below from which an affirmative answer to Problem B follows easily.

Analogously, one can study the corresponding problem about finite automata. A (generalized) finite automaton (in short, an nf automaton) $\mathscr{A}$ is a quintuple, $\langle\Sigma,D,Q,\delta,S,F\rangle$, where $D\subseteq\Sigma^*$ is a finite set called the input domain, Q is a finite set of states, $\delta:Q\times D\rightarrow 2^Q$ is a transition function, and $S\subseteq Q$ and $F\subseteq Q$ the sets of initial and final states, respectively. δ is extended to $\delta:Q\times\Sigma^*\rightarrow 2^Q$ and $\delta:2^Q\times\Sigma^*\rightarrow 2^Q$ in the standard way. $L(\mathscr{A})$ denotes the language accepted by $\mathscr{A}$, $\{w\in\Sigma^*\mid\delta(S,w)\cap F\neq\emptyset\}$, and we define $s(\mathscr{A})$ by $s(\mathscr{A})=|Q|$, where $\emptyset$ denotes the empty set.

For any regular language R, we define $s(R)$ by $s(R)=\min\{s(\mathscr{A})\mid\mathscr{A}$ is an nf automaton accepting $R\}$. An nf automaton $\mathscr{A}$ is R-state-minimal if $L(\mathscr{A})=R$ and $s(\mathscr{A})=s(R)$. In [2], an affirmative answer is presented to the following problem.

Problem C. Is it decidable to determine $s(R)$ for any given regular language R?

Analogous to the notion of $n(R_1,R_2)$, one can define the following. Let R_1 and R_2 be two regular languages. Define $s(R_1,R_2)$ by $s(R_1,R_2)=\min\{s(R)\mid R$ is a regular language with $R_1\subseteq R\subseteq R_2\}$ if $R_1\subseteq R_2$, and $s(R_1,R_2)=\infty$ otherwise. An nf automaton $\mathscr{A}$ is (R_1,R_2)-state-minimal if $R_1\subseteq L(\mathscr{A})\subseteq R_2$ and $s(\mathscr{A})=s(R_1,R_2)$.

Problem D. Is it decidable to determine $s(R_1,R_2)$ for any given two regular languages R_1,R_2?

Clearly, Problems A and C have certain similar characters, and so do Problems B and D. The following two facts can be seen easily.

Fact 1 (Hashiguchi [2]). *Let R be any regular language.*
(1) $n(R)\leqslant s(R)$ *and* $s(R)\leqslant n(R)+1$.
(2) *If R is finite, then* $n(R)=1$.
(3) *If R is finite, then*
  (3.1) $s(R)=1$ *if* $R=\emptyset$ *or* $\{\lambda\}$, *where* λ *is the null word.*
  (3.2) $s(R)=2$ *otherwise.*

Fact 2. *Let R_1 and R_2 be two regular languages with $R_1 \subseteq R_2$.*
(1) $n(R_1,R_2) \leqslant s(R_1,R_2)$ *and* $s(R_1,R_2) \leqslant n(R_1,R_2)+1$.
(2) *If R_1 is finite, then* $n(R_1,R_2)=n(R_1)=1$.
(3) *If R_1 is finite, then*
  (3.1) $s(R_1,R_2)=1$ *if* $R_1=\emptyset$ *or* $R_1=\{\lambda\}$.
  (3.2) $s(R_1,R_2)=2$ *otherwise*.
(4) *If* $R_1=R_2$, *then* $n(R_1,R_2)=n(R_1)$ *and* $s(R_1,R_2)=s(R_1)$.

In Section 2, we shall present outlines of two algorithms for determining $n(R_1,R_2)$ and $s(R_1,R_2)$, respectively, for any given two regular languages R_1,R_2. To do this, we need Theorems 2.1 and 2.2 below which can be proved by similar ideas as the ones used for proving Theorems 1 and 2 in [2]. For readability, we shall present a detailed proof for Theorem 2.1 (including the proof of Lemma 2.3 which appears in [2]) in this note.

2. Main results

Firstly, we shall present several definitions.

Definition 2.1. Let k,m be integers with $0 \leqslant k \leqslant m$. An rl grammar $G=\langle V,\Sigma,P,S\rangle$ is in (k,m)-form if the following hold:
(1) For any $S \to u \in P$ with $u \in \Sigma^*$, $|u| \leqslant m$.
(2) For any $S \to uB \in P$ with $u \in \Sigma^*$ and $B \in V$, $k \leqslant |u| \leqslant m$.
(3) For any $A \to uB \in P$ or $A \to v \in P$ with $A \in V-\{S\}$, $B \in V$ and $u,v \in \Sigma^*$, $k \leqslant |u|$, $|v| \leqslant m$ holds.

Definition 2.2. (1) For any rl grammar $G=\langle V,\Sigma,P,S\rangle$, $\mu(G)$ and $v(G)$ are defined by

$$\mu(G) = \{w \in \Sigma^* \mid A \to wB \in P \text{ or } A \to w \in P \text{ for some } A,B \in V\},$$

$$v(G) = \max\{|w| \mid w \in \mu(G)\}.$$

(2) For any two regular languages R_1 and R_2 with $R_1 \subseteq R_2$, $v(R_1,R_2)$ is defined by $v(R_1,R_2)=\min\{v(G) \mid G$ is an (R_1,R_2)-nonterminal-minimal rl grammar$\}$.

Lemma 2.3. *For any rl grammar $G=\langle V,\Sigma,P,S\rangle$ and any integer $k \geqslant 0$, there exists an rl grammar G' in $(k,v(G)+2k)$-form such that $L(G)=L(G')$ and $n(G')=n(G)$.*

Proof. When $k=0$, the assertion is trivial. Let $k \geqslant 1$. We define $G'=\langle V,\Sigma,P',S\rangle$ as follows:
$P'=P_0 \cup P_1 \cup P_2$, where $P_0=\{S \to u \mid u \in \Sigma^*, S \stackrel{*}{\Rightarrow}_G u$ and $|u| \leqslant 2k+v(G)\}$, $P_1=\{A \to uB \mid A,B \in V, u \in \Sigma^*, A \stackrel{*}{\Rightarrow}_G uB$ and $k \leqslant |u| \leqslant 2k+v(G)\}$, and $P_2=\{A \to u \mid A \in V-\{S\}, u \in \Sigma^*, A \stackrel{*}{\Rightarrow}_G u$ and $k \leqslant |u| \leqslant 2k+v(G)\}$.

It suffices to prove that $L(G)=L(G')$. Clearly, $L(G')\subseteq L(G)$. Conversely, consider any $w\in L(G)$. If $|w|\leqslant 2k+v(G)$, then clearly $w\in L(G')$. Otherwise, there exists a derivation of w such that $S\Rightarrow_G u_1A_1\Rightarrow_G \cdots \Rightarrow_G u_1u_2\cdots u_{n-1}A_{n-1}\Rightarrow_G u_1u_2\cdots u_n=w$ with $n\geqslant 2$. Since $|u_i|\leqslant v(G)$ for all $1\leqslant i\leqslant n$, one can see that there exists a derivation of the form $S\overset{*}{\Rightarrow}_G v_1B_1\overset{*}{\Rightarrow}_G v_1v_2B_2\overset{*}{\Rightarrow}_G \cdots \overset{*}{\Rightarrow}_G v_1v_2\cdots v_{m-1}B_{m-1}\overset{*}{\Rightarrow}_G v_1v_2\cdots v_m=w$ such that $1\leqslant m\leqslant n, k\leqslant|v_j|\leqslant k+v(G)$ for all $1\leqslant j\leqslant m-1$, and $k\leqslant|v_m|\leqslant 2k+v(G)$. These imply $w\in L(G')$. □

Now consider any two regular languages R_1 and R_2 with $R_1\subseteq R_2$. For $i=1,2$, let M_i be the syntactic monoid of R_i and β_i be the canonical morphism from Σ^* onto M_i. Thus, for any $u,v\in\Sigma^*,\beta_i(u)=\beta_i(v)$ iff (for any $x,y\in\Sigma^*, xuy\in R_i$ iff $xvy\in R_i$). Now we define a congruence relation $\equiv$ over Σ^* as follows:

For any $u,v\in\Sigma^*$

$u\equiv v\Leftrightarrow(\beta_1(u)=\beta_1(v)$ and $\beta_2(u)=\beta_2(v))$.

Let M denote the quotient monoid $\Sigma^*/\equiv$ and β be the canonical morphism from Σ^* onto M. Thus, it holds that for any $u,v\in\Sigma^*$

$\beta(u)=\beta(v)\Leftrightarrow(\beta_1(u)=\beta_1(v)$ and $\beta_2(u)=\beta_2(v))$.

Let m denote the cardinality of M. The following lemma can be proved as Lemma 5.1 in [1].

Lemma 2.4. *For any $w\in\Sigma^+$ of length $\geqslant m(m+2)$, there exist $x,y,z\in\Sigma^+$ such that $w=xyz,\beta(x)=\beta(xy)$ and $\beta(yz)=\beta(z)$.*

The following theorem is one of the main results of this note.

Theorem 2.1. *Let R_1 and R_2 be two regular languages with $R_1\subseteq R_2$ and m be as above. Then there exists an (R_1,R_2)-nonterminal-minimal rl grammar G such that $R_1\subseteq L(G)\subseteq R_2$ and $v(G)\leqslant 2m(m+2)(4m(m+2)+3)$.*

Proof. Let $G=\langle V,\Sigma,P,S\rangle$ be an (R_1,R_2)-nonterminal-minimal rl grammar such that $R_1\subseteq L(G)\subseteq R_2$, $n(G)=n(R_1,R_2)$, and $v(G)=v(R_1,R_2)$. Assume that $v(G)>2m(m+2)(4m(m+2)+3)$. Then R_1 is not finite since $v(G)>m$ implies that there exists $w\in R_1$ with $|w|>m$. (If R_1 is finite, then $n(R_1,R_2)=1$ and it suffices to let G be such that $L(G)=R_1$.) We shall derive a contradiction. By Lemma 2.3, there exists a nonterminal-minimal rl grammar $G'=\langle V,\Sigma,P',S\rangle$ in $(2m(m+2),v(G)+4m(m+2))$-form with $L(G')=L(G)$. We define a set X by

$X=\{(x,y,z)\mid x,y,z\in\Sigma^+,|xyz|=m(m+2),\beta(xy)=\beta(x)$ and $\beta(yz)=\beta(z)\}$.

Now consider any $w\in R_1$ with $|w|\geqslant 2m(m+2)$. w has a decomposition of the form

$$w=x_1y_1z_1x_2y_2z_2\cdots x_py_pz_pw_0,$$

where for $1\leqslant i\leqslant p$, $(x_i,y_i,z_i)\in X$ and $|w_0|<m(m+2)$. Then the word $w'=x_1y_1y_1z_1x_2y_2y_2z_2\cdots x_py_py_pz_pw_0$ is also in R_1, and w' has a derivation in G'. In this derivation of w', if we delete each one of $y_i(1\leqslant i\leqslant p)$, then we could obtain a derivation of w

in another rl grammar G''. Thus, we shall construct an rl grammar $G'' = \langle V, \Sigma, P'', S\rangle$ in the following way.

For each $w \in \mu(G')$, we define a set $Y(w)$ (by which w will be replaced in P' to obtain P''), as follows:

$$Y(w) = C_1(w) \cup C_2(w),$$

where

(1) $C_1(w) = \{v \in \Sigma^* \mid |v| \leqslant 2m(m+2) \text{ and } \beta(v) = \beta(w)\}$.

(2) $C_2(w) = \{v \in \Sigma^* \mid$ for some $k \geqslant 1, u_0, u_1 \in \Sigma^*$ with length $< m(m+2)$, and $(x_i, y_i, z_i) \in X$ for $1 \leqslant i \leqslant k$, v satisfies one of the following, (2.1)–(2.4)$\}$.

(2.1) $w = u_0x_1y_1y_1z_1 \cdots x_ky_ky_kz_ku_1$ and $v = u_0x_1y_1z_1 \cdots x_ky_kz_ku_1$,

(2.2) $w = u_0y_1z_1x_2y_2y_2z_2 \cdots x_ky_ky_kz_ku_1$ and $v = u_0y_1z_1x_2y_2z_2 \cdots x_ky_kz_ku_1$ or $v = u_0z_1x_2y_2z_2 \cdots x_ky_kz_ku_1$,

(2.3) $w = u_0x_1y_1y_1z_1 \cdots x_{k-1}y_{k-1}y_{k-1}z_{k-1}x_ky_ku_1$ and $v = u_0x_1y_1z_1 \cdots x_{k-1}y_{k-1}z_{k-1}x_ky_ku_1$ or $v = u_0x_1y_1z_1 \cdots x_{k-1}y_{k-1}z_{k-1}x_ku_1$,

(2.4) $w = u_0y_1z_1x_2y_2y_2z_2 \cdots x_{k-1}y_{k-1}y_{k-1}z_{k-1}x_ky_ku_1$ and $v = u_0y_1z_1x_2y_2z_2 \cdots x_{k-1}y_{k-1}z_{k-1}x_ky_ku_1$ or $v = u_0z_1x_2y_2z_2 \cdots x_{k-1}y_{k-1}z_{k-1}x_ky_ku_1$ or $v = u_0y_1z_1x_2y_2z_2 \cdots x_{k-1}y_{k-1}z_{k-1}x_ku_1$ or $v = u_0z_1x_2y_2z_2 \cdots x_{k-1}y_{k-1}z_{k-1}x_ku_1$.

One can see the following.

(3) $C_2(w)$ may be empty, but $C_1(w)$ is not empty, and so $Y(w)$ is not empty.

(4) For any $v \in Y(w)$, either $|w|, |v| < v(G)$, or else $2m(m+2)(4m(m+2)+3) \leqslant v(G) \leqslant |w| \leqslant v(G) + 4m(m+2)$ and

$$|v| \leqslant |w| - \frac{|w| - 4m(m+2)}{2m(m+2)}$$
$$\leqslant v(G) + 4m(m+2) - \frac{2m(m+2)(4m(m+2)+3) - 4m(m+2)}{2m(m+2)} < v(G).$$

We define P'' by

$$P'' = \{A \to v \mid A \to w \in P' \text{ for some } A \in V \text{ and } w \in \Sigma^*, \text{ and } v \in Y(w)\}$$
$$\cup \{A \to vB \mid A \to wB \in P' \text{ for some } A, B \in V \text{ and } w \in \Sigma^*, \text{ and } v \in Y(w)\}.$$

It is clear that $v(G'') < v(G)$ and $n(G'') = n(G)$. To derive a contradiction, it suffices to show $R_1 \subseteq L(G'') \subseteq R_2$. Due to the definitions of β and P'', it is easy to see that $L(G'') \subseteq R_2$. To show $R_1 \subseteq L(G'')$, consider any $w \in R_1$. If $|w| \leqslant 2m(m+2)$, then clearly $w \in L(G'')$. Otherwise, we consider a decomposition of w, $w = x_1y_1z_1 \cdots x_py_pz_pw_0$ and the corresponding w' as above, and can see that $w \in L(G'')$. □

Now we shall present an algorithm for Problem B briefly. Assume that two regular languages R_1 and R_2 are given. We may assume without loss of generality that we are given two rl grammars G_1 and G_2 which generate R_1 and R_2, respectively. Decide whether or not $R_1 \subseteq R_2$. If $R_2 - R_1 \neq \emptyset$, then $n(R_1, R_2) = \infty$. Otherwise, we construct the syntactic monoids M_1 and M_2 of R_1 and R_2, respectively, the congruence monoid $M = \Sigma^*/\equiv$ and the canonical morphism β as above. Put $m = |M|$, and construct the following finite family $F(R_1, R_2)$ of rl grammars:

$F(R_1, R_2) = \{G \mid G$ is an rl grammar, $R_1 \subseteq L(G) \subseteq R_2$, $v(G) \leqslant 2m(m+2)(4m(m+2) + 3)$ and $n(G) \leqslant \min\{n(G_1), n(G_2)\}\}$.

Then $n(R_1, R_2)$ can be determined by

$$n(R_1, R_2) = \min\{n(G) \mid G \in F(R_1, R_2)\}.$$

Now we shall study Problem D.

Definition 2.5. Let k, m be integers with $0 \leqslant k \leqslant m$. An nf automaton $\mathscr{A} = \langle \Sigma, D, Q, \delta, S, F \rangle$ is in (k, m)-form if for any $(q, w) \in Q \times D$, the following (1)–(3) hold.
(1) If $q \in S$ and $\delta(q, w) \subseteq F$, then $|w| \leqslant m$.
(2) If $q \in S$ and $\delta(q, w) - F \neq \emptyset$, then $k \leqslant |w| \leqslant m$.
(3) If $q \in Q - S$ and $\delta(q, w) \neq \emptyset$, then $k \leqslant |w| \leqslant m$.

Definition 2.6. (1) For any nf automaton $\mathscr{A} = \langle \Sigma, D, Q, \delta, S, F \rangle$, $D(\mathscr{A})$ is defined by

$$D(\mathscr{A}) = \max\{|w| \mid w \in D\}.$$

(2) For any two regular languages R_1 and R_2 with $R_1 \subseteq R_2$, $D(R_1, R_2)$ is defined by $D(R_1, R_2) = \min\{D(\mathscr{A}) \mid \mathscr{A}$ is an (R_1, R_2)-state-minimal nf automaton with $R_1 \subseteq L(\mathscr{A}) \subseteq R_2\}$.

Lemma 2.7. *For any nf automaton $\mathscr{A} = \langle \Sigma, D, Q, S, F \rangle$ and any integer $k \geqslant 0$, there exists an nf automaton $\mathscr{A}'$ in $(k, D(\mathscr{A}) + 2k)$-form such that $L(\mathscr{A}') = L(\mathscr{A})$ and $s(\mathscr{A}') = s(\mathscr{A})$.*

Proof. When $k = 0$, the assertion is trivial. Let $k \geqslant 1$. We define $\mathscr{A}' = \langle \Sigma, D', Q, \delta', S, F \rangle$ as follows:
(1) $D' = D_0 \cup D_1$, where $D_0 = \{w \in L(\mathscr{A}) \mid |w| < k\}$ and $D_1 = \{w \in \Sigma^+ \mid k \leqslant |w| \leqslant 2k + D(\mathscr{A})\}$,
(2) $\delta'(p, w) = \delta(p, w) \cap F$ for any $p \in S$ and $w \in D_0$,
(3) $\delta'(q, w) = \delta(q, w)$ for any $q \in Q$ and $w \in D_1$,
(4) For any other $(q, w) \in D'$, $\delta'(q, w) = \emptyset$.
Here, δ is extended to $\delta : Q \times \Sigma^* \to 2^Q$ in the standard way. Now it suffices to show $L(\mathscr{A}') = L(\mathscr{A})$. Clearly, $L(\mathscr{A}') \subseteq L(\mathscr{A})$. Conversely, consider any $w \in L(\mathscr{A})$. If $|w| \leqslant 2k + D(\mathscr{A})$, then clearly $w \in L(\mathscr{A}')$. Otherwise, there exists an accepting transition sequence of w in $\mathscr{A}$, $p \to_{v_1} q_1 \to_{v_2} q_2 \to \cdots \to q_{n-1} \to_{v_n} q_n$, where $p \in S, n \geqslant 2, q_i \in Q$ for all i, $q_n \in F$ and $w = v_1 \cdots v_n$. As in the proof of Lemma 2.31, one can see that there exists an accepting transition sequence of w in $\mathscr{A}'$ and so $w \in L(\mathscr{A}')$. $\square$

One can prove the following theorem similarly as Theorem 2.1.

Theorem 2.2. *Let R_1 and R_2 be two regular languages and m be the integer as in Theorem 2.1. Then, there exists an (R_1, R_2)-state-minimal nf automaton $\mathscr{A}$ such that $R_1 \subseteq L(\mathscr{A}) \subseteq R_2$ and $D(\mathscr{A}) \leqslant 2m(m+2)(4m(m+2)+3)$.*

As in the case of Problem B, one can construct an algorithm for determining $n(R_1, R_2)$ for any given two regular languages R_1 and R_2. This provides an affirmative answer to Problem D.

In [2], the following problem is presented as an open problem.

Is it decidable to determine the smallest number of star operators which are sufficient for denoting a given regular language in a (restricted) regular expression?

As far as the author knows, this problem still remains open.

References

[1] K. Hashiguchi, Regular languages of star height one, Inform. Control 53 (1982) 199–210.

[2] K. Hashiguchi, Algorithms for determining the smallest number of nonterminals (states) sufficient for generating (accepting) a regular language, Proc. 18th Internat. Coll. on Automata, Languages and Programming, 1991, Lecture Notes in Computer Science, Vol. 510, Springer, Berlin, 1991, pp. 641–648.

[3] O.H. Ibarra, The unsolvability of the equivalence problem for ε-free NGSM's with unary input (output) alphabet and applications, SIAM J. Comput. 7 (1973) 524–532.