Decidability results concerning tree transducers II

By Z. ÉSIK

1. Introduction

Let $\tau \subseteq T_F \times T_G$ be an arbitrary tree transformation induced by a top-down or bottom-up tree transducer **A**. It is said that **A** preserves regularity if $\tau(R)$ is a regular forest for each regular forest $R \subseteq T_F$. It is natural to raise the question whether the regularity preserving property of tree transducers is decidable or not. This question was positively answered for bottom-up transducers in [4]. Even more, it was shown that a bottom-up transducer preserves regularity if and only if it is equivalent to a linear bottom-up transducer. Concerning top-down transducers we have quiet different results. Although every linear top-down transducer preserves regularity as linear top-down tree transformations form a (proper) subclass of linear bottom-up transformations (cf. [2]), there are deterministic regularity preserving top-down tree transducers having no linear bottom-up equivalent. Another distinction lies in the fact that there is no algorithm which can decide the regularity preserving property of top-down transducers (cf. Theorem 2). However, restricting ourselves to deterministic top-down transducers we obtain positive result (cf. Theorem 1).

The notations will be used in accordance with [1]. Recall that a top-down tree transducer $\mathbf{A}=(F, A, G, A_0, \Sigma)$ is called uniform if each rewriting rule in Σ is of the form $af \to q(a_1x_1, \ldots, a_nx_n)$ where $n \geqq 0, f \in F_n$, $a, a_1, \ldots, a_n \in A$ and $q \in T_{G,n}$. In addition, if q is always linear (cf. [2]) then **A** is called linear. These concepts extend to top-down tree transducers with regular look-ahead, as well. Furthermore, one-state top-down tree transducers and their induced transformations will be called homomorphisms. If **A** is a homomorphism then we omit the single state in the presentation of Σ.

2. Deterministic top-down transducers

Let $\mathbf{A}=(F, A, G, a_0, \Sigma)$ be an arbitrary deterministic top-down transducer kept fixed in this section. Put $\tau=\tau_{\mathbf{A}}$. If there exist

$$n_1, n_2, m_1, m_2 \geqq 0, \quad \mathbf{a}\in A^{n_1}, \quad \mathbf{b}\in A^{m_1}, \quad \mathbf{c}\in A^{n_2}, \quad \mathbf{d}\in A^{m_2}, \quad p_0, p_1\in \hat{T}_{F,1}$$

$$p_2\in T_F, \quad q_0\in \hat{T}_{G,n_1+m_1}, \quad \mathbf{q}_1\in \hat{T}^{n_1}_{G,n_2}, \quad \mathbf{r}_1\in \hat{T}^{m_1}_{G,m_2}, \quad \mathbf{q}_2\in T_G^{n_2}, \quad \mathbf{r}_2\in T_G^{m_2}$$

such that we have

$$a_0 p_0 \overset{*}{\Rightarrow} q_0(\mathbf{a}\mathbf{x}_1^{n_1}, \mathbf{b}\mathbf{x}_1^{m_1}),$$

$$\mathbf{a}_1 \mathbf{p}_1^{n_1} \overset{*}{\Rightarrow} \mathbf{q}_1(\mathbf{c}\mathbf{x}_1^{n_2}), \quad \mathbf{b}\mathbf{x}_1^{m_1} \overset{*}{\Rightarrow} \mathbf{r}_1(\mathbf{d}\mathbf{x}_1^{m_2}),$$

$$\mathbf{c}\mathbf{p}_2^{n_2} \overset{*}{\Rightarrow} \mathbf{q}_2, \quad \mathbf{d}\mathbf{p}_2^{m_2} \overset{*}{\Rightarrow} \mathbf{r}_2,$$

$$\{a_i|i\in[n_1]\} = \{c_i|i\in[n_2]\}, \quad \{b_i|i\in[m_1]\} = \{d_i|i\in[m_2]\},$$

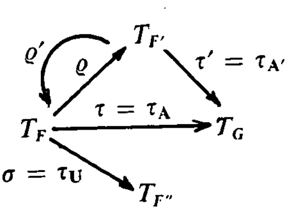
and both $\mathbf{q}_1$ and $\mathbf{r}_1$ contain an occurence of a symbol from G then we say that $\mathbf{A}$ satisfies condition $(*)$. Observe that our conditions imply that $n_i, m_i>0$ $(i=1, 2)$.

We are going to prove that $\mathbf{A}$ preserves regularity if and only if $(*)$ is not satisfied by $\mathbf{A}$. The necessity of this statement can be proved easily.

Lemma 1. If $\mathbf{A}$ preserves regularity then $\mathbf{A}$ does not satisfy condition $(*)$.

Proof. Assume that $\mathbf{A}$ satisfies condition $(*)$. Then, using the notations of the definition above, set $R=\{p_0(\underbrace{p_1(\dots(p_1}_{n\text{-times}}(p_2))\dots))|n\geqq 0\}$. R is regular and $\tau(R)$ consists of trees $q_0(\mathbf{r}_n, \mathbf{s}_n)$ $(n\geqq 0, \mathbf{r}_n\in T_G^{n_1}, \mathbf{s}_n\in T_G^{m_1})$ with the property that $n<\mathrm{rn}(\mathbf{r}_n)<\mathrm{rn}(\mathbf{r}_{n+1})$, $n<\mathrm{rn}(\mathbf{s}_n)<\mathrm{rn}(\mathbf{s}_{n+1})$[1]. Suppose that $\tau(R)$ is recognizable by a deterministic tree automaton $\mathbf{D}=(G, D, D_0)$. Let $n>m_1(1+\nu(G)+\dots+\nu(G)^{|D|-1})$ be an arbitrary fixed integer. As $(q_0(\mathbf{r}_n, \mathbf{s}_n))_{\mathbf{D}}\in D_0$ also there is a vector of trees $\mathbf{s}\in T_G^{m_1}$ with $\mathrm{dp}(\mathbf{s})<|D|$ and $(q_0(\mathbf{r}_n, \mathbf{s}))_{\mathbf{D}}\in D_0$. However, as $\mathrm{dp}(\mathbf{s})<|D|$ we obtain that $\mathrm{rn}(\mathbf{s})\leqq \leqq m_1(1+\nu(G)+\dots+\nu(G)^{|D|-1})$. This contradicts $\tau(R)=T(\mathbf{D})$. Therefore, $\tau(R)$ is not regular, as was to be proved.

To prove the converse of Lemma 1 first we show that $\tau(\mathrm{dom}\,\tau)$ is regular if $\mathbf{A}$ does not satisfy $(*)$. This will be carried out by constructing a linear deterministic top-down tree transducer with regular look-ahead such that $\tau(\mathrm{dom}\,\tau)= =\tau_{\mathbf{A}'}(\mathrm{dom}\,\tau_{\mathbf{A}'})$. The construction of $\mathbf{A}'$ will be made by the help of other tree transformations. Thus, we shall have the transformations indicated by the figure below:



[1] $\mathrm{rn}(\mathbf{r}_n)=\mathrm{rn}(r_{n1})+\dots+\mathrm{rn}(r_{nn_1})$, $\mathrm{rn}(\mathbf{s}_n)$ is similarly defined.

We begin with the definition of F''. First let $\bar{F}=\bigcup_{n\geqq 0} \bar{F}_n$, $\bar{F}_n=\{(f, C, \varphi, \psi)|f\in F_n,$ $C\subseteqq B$, $\varphi\colon B\to P(A)$, $\psi\colon A\rightarrowtail A$ for a subset $B\subseteqq A\}$, i.e. φ is a mapping of B into the power-set of A and ψ is a partial function on A. Now the type F'' is defined by $F''_n=F_n\cup\bar{F}_n$ $(n\geqq 0)$.

The $\bar{F}$-depth $(\overline{\mathrm{dp}}(p))$ and $\bar{F}$-width $(\overline{\mathrm{wd}}(p))$ of a tree $p\in T_{F''}$ are defined by

$$\overline{\mathrm{dp}}(p)=0,\quad \overline{\mathrm{wd}}(p)=\overline{\mathrm{wd}}_0(p)=0 \quad\text{if}\quad p\in F_0,$$

$$\overline{\mathrm{dp}}(p)=1,\quad \overline{\mathrm{wd}}(p)=\overline{\mathrm{wd}}_0(p)=1 \quad\text{if}\quad p\in \bar{F}_0,$$

$$\overline{\mathrm{dp}}(p)=\max\{\overline{\mathrm{dp}}(p_i)|i\in[n]\},\quad \overline{\mathrm{wd}}(p)=\max\left\{\sum_{i=1}^{n}\overline{\mathrm{wd}}_0(p_i),\quad \overline{\mathrm{wd}}(p_i)|i\in[n]\right\},$$

$$\overline{\mathrm{wd}}_0(p)=\sum_{i=1}^{n}\overline{\mathrm{wd}}_0(p_i)\quad\text{if}\quad p=f(p_1,\dots,p_n)\ \text{with}\ n>0,\ f\in F_n,$$

$$p_1,\dots,p_n\in T_{F''},$$

$$\overline{\mathrm{dp}}(p)=1+\max\{\overline{\mathrm{dp}}(p_i)|i\in[n]\},\quad \overline{\mathrm{wd}}(p)=\max\{1, \overline{\mathrm{wd}}(p_i)|i\in[n]\},$$

$$\overline{\mathrm{wd}}_0(p)=1\quad\text{if}\quad p=f(p_1,\dots,p_n)\ \text{where}\ n>0,\ f\in\bar{F}_n,$$

$$p_1,\dots,p_n\in T_{F''}.$$

If n, m are given nonnegative integers then $T_{(n,m)}$ denotes the set of all trees $p\in T_{F''}$ with $\overline{\mathrm{dp}}(p)<n$ and $\overline{\mathrm{wd}}(p)\leqq m$.

We shall frequently use an equivalence relation denoted by $\sim$ on $T_{F''}$. Given $p, q\in T_{F''}$, $p\sim q$ if and only if one of the following three conditions holds:

(i) $p, q\in T_F$,

(ii) $p=f(p_1,\dots,p_n)$, $q=f(q_1,\dots,q_n)$ with $n\geqq 0$, $f\in\bar{F}_n$, $p_i, q_i\in T_{F''}$ and $p_i\sim q_i$ $(i\in[n])$,

(iii) $p=p_0(p_1,\dots,p_n)$, $q=q_0(q_1,\dots,q_n)$ with $n>0$, $p_0, q_0\in\tilde{T}_{F,n}$, $p_i, q_i\in T_{F''}$, $p_i\sim q_i$, $\mathrm{rt}(p_i)$, $\mathrm{rt}(q_i)\in\bar{F}$ $(i\in[n])$.

If $p\in T_{F''}$ then $[p]$ denotes the block containing p under the partition induced by $\sim$.

The next statement can be proved in an easy way.

Lemma 2. $[p]$ is a regular forest for any $p\in T_{F''}$.

Now we introduce the transducer **U**. $\mathbf{U}=(F, U, F'', u_0, \Sigma'')$ where

$$U=\{(B, B', C, \varphi, \psi)|B\subseteqq A,\ B'\subseteqq B,\ C\subseteqq B,\ \varphi\colon B\to P(A),\ \psi\colon A\rightarrowtail A\},$$

$a_0=(\{a_0\}, \emptyset, \emptyset, \varphi, \psi)$ with $\varphi(a_0)=\{a_0\}$ and $\psi(a)=b$ if and only if $a=b=a_0$. Σ'' is determined as follows.

Let $u=(B, B', C, \varphi, \psi)$ be an arbitrary element of U, $f\in F_n$ $(n\geqq 0)$. Assume that in Σ there is a rule with left side af for any $a\in\bigcup\varphi(B)$. That is, $\bigcup\varphi(B)=$ $=\{a_1,\dots,a_l, a_{11},\dots,a_{l_1 1},\dots,a_{1n},\dots,a_{l_n n}\}$ $(l, l_1,\dots,l_n\geqq 0)$ and

$$a_i f\to q_i(\mathbf{b}_{i1}\mathbf{x}_1^{k_{i1}},\dots,\mathbf{b}_{in}\mathbf{x}_n^{k_{in}})\in\Sigma\quad (i\in[l]),$$

$$a_{ij}f\to c_{ij}x_j\in\Sigma\quad (i\in[l_j],\ j\in[n]),$$

where $k_{ij}\geqq 0$ $(i\in[l], j\in[n])$, $q_i\in\hat{T}_{G,k_i}-X$, $k_i=\sum_{j=1}^{n} k_{ij}$ $(i\in[l])$, $\mathbf{b}_{ij}\in A^{k_{ij}}$ $(i\in[l], j\in[n])$, $c_{ij}\in A$ $(i\in[l_j], j\in[n])$.

Then Σ'' is the smallest set of top-down rewriting rules satisfying (i) and (ii) below.

$$\text{(i) If } \left|\{a\in B|\varphi(a)\cap\{a_1, \dots, a_l\}\neq\emptyset\}\right|\geqq 2 \quad\text{or}$$

$$\left|\{a\in B'|\varphi(a)\cap\{a_1, \dots, a_l\}\neq\emptyset\}\right|\geqq 1 \quad\text{or}$$

$$\left|\{a\in B-C|\varphi(a)\cap\{a_1, \dots, a_l\}\neq\emptyset\}\right|\geqq 1 \quad\text{or}$$

$$|C|\geqq 2 \quad\text{and}\quad \left|\{a\in B|\varphi(a)\cap\{a_1, \dots, a_l\}\neq\emptyset\}\right|\geqq 1$$

and

$$u_i=(B, B', C', \varphi_i, \psi_i)\quad (i\in[n]),$$

$$C'=\{a\in B|\varphi(a)\cap\{a_1, \dots, a_l\}\neq\emptyset\},$$

$$\varphi_i(a)=\bigcup(B_{ji}|a_j\in\varphi(a))\cup\{c_{ji}|a_{ji}\in\varphi(a)\}\quad (a\in B,\ i\in[n]),$$

where B_{ji} denotes the set of components of the vector $\mathbf{b}_{ji}$ $(j\in[l], i\in[n])$, $\psi_i(a)=b$ if and only if $a=b\in\bigcup\varphi_i(B)$ $(i\in[n])$, $\bar{f}=(f, C', \varphi, \psi)$ then

$$uf\to\bar{f}(u_1x_1, \dots, u_nx_n)\in\Sigma''.$$

(ii) If not (i), i.e. $l=0$ or
$l=1$, $C=\{a_1\}$ and $a_1\notin B'$
and for each $i\in[n]$
$u_i=(B, B', C, \varphi_i, \psi_i)$,
φ_i is the same as in the previous case,
$\psi_i=\psi\circ\psi_i'$ with $\psi_i'(a)=b$ if and only if $a=a_{ji}$, $b=c_{ji}$ $(j\in[l_i])$,
then
$uf\to f(u_1x_1, \dots, u_nx_n)\in\Sigma''$.

Observe that $\mathbf{U}$ is a deterministic top-down relabeling. The following properties of $\mathbf{U}$ will be used without any reference. First, if $up\overset{*}{\underset{\mathbf{U}}{\Rightarrow}}q(\mathbf{v}(x_1, \dots, x_n))$ $(n\geqq 0, u\in U, v\in U^n)$ and $p, q\in\tilde{T}_{F,n}$ then $p=q$. Secondly, let $\mathbf{a}\in A^k$ $(k\geqq 0)$ be arbitrary and identify $\mathbf{a}$ with the state $u=(B, B', \emptyset, \varphi, \psi)$ where $B=\{a_i|i\in[k]\}$, $B'=\{a_i|i\in[k],\ \exists j\in[k]\ i\neq j,\ a_i=a_j\}$, $\varphi(a)=\{a\}$ if $a\in B$ and $\psi(a)=b$ if and only if $a=b\in B$. Denote by $\sigma_{\mathbf{a}}$ the transformation $\tau_{\mathbf{U}(u)}$ and similarly, put $\tau_{a_i}=\tau_{\mathbf{A}(a_i)}$ $(i\in[k])$. Then, for any $p\in T_F$, $p\in\operatorname{dom}\sigma_{\mathbf{a}}$ if and only if $p\in\bigcap_{i=1}^{k}\operatorname{dom}\tau_{a_i}$.

In the next few lemmata we shall point to further connections between $\mathbf{A}$ and $\mathbf{U}$.

Lemma 3. If $\mathbf{A}$ does not satisfy condition $(*)$ then $\overline{\mathrm{dp}}(\sigma_{\mathbf{a}}(p))<2|A|^2\|A\|^2$ holds for any $\mathbf{a}\in A^k$ $(k\geqq 0)$ and $p\in T_F$ provided that $\sigma_{\mathbf{a}}(p)$ is defined and there exist trees $r\in\hat{T}_{F,1}$ and $r'\in\hat{T}_{G,k}$ with $a_0r\overset{*}{\underset{\mathbf{A}}{\Rightarrow}}r'(\mathbf{a}\mathbf{x}_1^k)$.

Proof. Let $L=|A|^2\|A\|^2$ and suppose that $a_0r\overset{*}{\underset{\mathbf{A}}{\Rightarrow}}r'(\mathbf{a}\mathbf{x}_1^k)$ and $\overline{\mathrm{dp}}(\sigma_a(p))\geqq 2L$. Then $k\geqq 2$ and there exist $p_0, \dots, p_{2L-1}\in\hat{T}_{F,1}$, $p_{2L}\in T_F$, $q_0, \dots, q_{2L-1}\in\hat{T}_{F'',1}$,

$q_{2L} \in T_{F''}$, $u_1, \ldots, u_{2L} \in U$ such that

$$p = p_0(\ldots(p_{2L})\ldots), \quad q = q_0(\ldots(q_{2L})\ldots),$$

$$\mathbf{a}p_0 \overset{*}{\underset{\mathbf{u}}{\Rightarrow}} q_0(u_1 x_1), \quad u_i p_i \overset{*}{\underset{\mathbf{u}}{\Rightarrow}} q_i(u_{i+1} x_1) \quad (i = 1, \ldots, 2L-1), \quad u_{2L} p_{2L} \overset{*}{\underset{\mathbf{u}}{\Rightarrow}} q_{2L},$$

furthermore, $\mathrm{rt}(q_i) \in \bar{F}$, say, $\mathrm{rt}(q_i) = (f_i, C_i, \varphi_i, \psi_i)$ $(i=1, \ldots, 2L)$. Let $D_1 = C_1 \cup C_2, \ldots, D_L = C_{2L-1} \cup C_{2L}$. It is not difficult to see by the definition of $\mathbf{U}$ that for any $i \in [L]$ there exist indices $j_i \neq k_i$ $(j_i, k_i \in [k])$ with $a_{j_i}, a_{k_i} \in D_i$. On the other hand, as $L = |A|^2 \|A\|^2$, there exist $i_1 < i_2$ $(i_1, i_2 \in [L])$ such that $a_{j_{i_1}} = a_{j_{i_2}}$, $a_{k_{i_1}} = a_{k_{i_2}}$, $S_{i_1} = S_{i_2}$ and $T_{i_1} = T_{i_2}$ where S_i and T_i are defined by $S_i = \varphi_{2i-1}(a_{j_i})$ and $T_i = \bigcup(\varphi_{2i-1}(a_j) | j \in [k], j \neq j_i)$. Without loss of generality we may take $j_{i_1} = j_{i_2} = 1$ and $k_{i_1} = k_{i_2} = 2$.

As $\sigma_{\mathbf{a}}(p)$ is defined also $\tau_{a_i}(p)$ is defined for any $i \in [k]$. Thus, if $r_1 = p_0(\ldots(p_{2i_1-2})\ldots)$, $r_2 = p_{2i_1-1}(\ldots(p_{2i_2-2})\ldots)$, $r_3 = p_{2i_2-1}(\ldots(p_{2L})\ldots)$ then the derivations

$$a_1 r_1 \overset{*}{\underset{\mathbf{A}}{\Rightarrow}} s_1(\mathbf{c}_1 \mathbf{x}_1^{n_1}), \quad (a_2, \ldots, a_k)\mathbf{r}_1^{k-1} \overset{*}{\underset{\mathbf{A}}{\Rightarrow}} \mathbf{t}_1(\mathbf{d}_1 \mathbf{x}_1^{m_1}),$$

$$\mathbf{c}_1 \mathbf{r}_2^{n_1} \overset{*}{\underset{\mathbf{A}}{\Rightarrow}} \mathbf{s}_2(\mathbf{c}_2 \mathbf{x}_1^{n_2}), \quad \mathbf{d}_1 \mathbf{r}_2^{m_1} \overset{*}{\underset{\mathbf{A}}{\Rightarrow}} \mathbf{t}_2(\mathbf{d}_2 \mathbf{x}_1^{m_2}),$$

$$\mathbf{c}_2 \mathbf{r}_3^{n_2} \overset{*}{\underset{\mathbf{A}}{\Rightarrow}} \mathbf{s}_3, \quad \mathbf{d}_2 \mathbf{r}_3^{m_2} \overset{*}{\underset{\mathbf{A}}{\Rightarrow}} \mathbf{t}_3$$

exist where $s_1 \in \hat{T}_{G,n_1}$, $\mathbf{t}_1 \in \hat{T}^{k-1}_{G,m_1}$, $\mathbf{s}_2 \in \hat{T}^{n_1}_{G,n_2}$, $\mathbf{t}_2 \in \hat{T}^{m_1}_{G,m_2}$, $\mathbf{s}_3 \in T^{n_2}_G$, $\mathbf{t}_3 \in T^{m_2}_G$ and $\mathbf{c}_i \in A^{n_i}$, $\mathbf{d}_i \in A^{m_i}$ $(i=1, 2)$.

Since $1, 2 \in D_{i_1}$ we have that both $\mathbf{s}_2$ and $\mathbf{t}_2$ contain an occurence of a symbol from G. Furthermore, as the sets $S_{i_1}, S_{i_2}, T_{i_1}$ and T_{i_2} coincide with the set of components of $\mathbf{c}_1, \mathbf{c}_2, \mathbf{d}_1$ and $\mathbf{d}_2$, respectively, it follows that $\mathbf{c}_1$ and $\mathbf{d}_1$ have the same set of components as $\mathbf{c}_2$ and $\mathbf{d}_2$.

By

$$a_0 r(r_1) \overset{*}{\underset{\mathbf{A}}{\Rightarrow}} r'(s_1(\mathbf{c}_1 \mathbf{x}_1^{n_1}), \quad \mathbf{t}_1(\mathbf{d}_1 \mathbf{x}_1^{m_1})),$$

$$\mathbf{c}_1 \mathbf{r}_2^{n_1} \overset{*}{\underset{\mathbf{A}}{\Rightarrow}} \mathbf{s}_2(\mathbf{c}_2 \mathbf{x}_1^{n_2}), \quad \mathbf{d}_1 \mathbf{r}_2^{m_1} \overset{*}{\underset{\mathbf{A}}{\Rightarrow}} \mathbf{t}_2(\mathbf{d}_2 \mathbf{x}_1^{m_2}),$$

$$\mathbf{c}_2 \mathbf{r}_3^{n_2} \overset{*}{\underset{\mathbf{A}}{\Rightarrow}} \mathbf{s}_3, \quad \mathbf{d}_2 \mathbf{r}_3^{m_2} \overset{*}{\underset{\mathbf{A}}{\Rightarrow}} \mathbf{t}_3$$

this yields that $\mathbf{A}$ satisfies condition $(*)$, which is a contradiction.

Lemma 4. Let $\mathbf{a} \in A^k$ $(k \geqq 0)$ be arbitrary. Put $B = \{a_i | i \in [k]\}$ and assume that

$$\mathbf{a}p_0 \overset{*}{\underset{\mathbf{u}}{\Rightarrow}} p_0(\mathbf{u}(x_1, \ldots, x_n)), \quad \mathbf{a}p_0' \overset{*}{\underset{\mathbf{u}}{\Rightarrow}} p_0'(\mathbf{u}'(x_1, \ldots, x_n)),$$

$$\mathbf{u}\mathbf{p} \overset{*}{\underset{\mathbf{u}}{\Rightarrow}} \mathbf{q}, \quad \mathbf{u}'\mathbf{p}' \overset{*}{\underset{\mathbf{u}}{\Rightarrow}} \mathbf{q}',$$

$$\mathrm{rt}(\mathbf{q}) = \mathrm{rt}(\mathbf{q}') \in \bar{F}^n$$

where $n \geqq 0$, $p_0, p_0' \in \tilde{T}_{F,n}$, $\mathbf{p}, \mathbf{p}' \in T_F^n$, $\mathbf{q}, \mathbf{q}' \in T_{F''}^n$, $\mathbf{u}, \mathbf{u}' \in U^n$.

Then $n \leqq |A|$ and $\tau_b(p_0(\mathbf{p})) = \tau_b(p_0'(\mathbf{p}))$ for any $b \in B$.

Proof. Suppose that $\mathrm{rt}(q_i)=(f_i, C_i, \varphi_i, \psi_i)$ $(i\in[n])$. It is not difficult to see by the definition of $\mathbf{U}$ that for any $i\in[n]$ there is a state $b\in B$ with $\psi_i(b)$ being defined and $bp_0 \overset{*}{\underset{\mathbf{A}}{\Rightarrow}} \psi_i(b)x_i$. Therefore, $n\leqq|B|$ and also $n\leqq|A|$. Similarly, for each $b\in B$ there is an integer $i\in[n]$ such that $\psi_i(b)$ is defined and $bp_0 \overset{*}{\underset{\mathbf{A}}{\Rightarrow}} \psi_i(b)x_i$, $bp_0' \overset{*}{\underset{\mathbf{A}}{\Rightarrow}} \psi_i(b)x_i$. From this $\tau_b(p_0(\mathbf{p}))=\tau_b(p_0'(\mathbf{p}))$ follows immediately.

Lemma 5. Let $\mathbf{a}\in A^k$ $(k>0)$ and define the set B as previously. Set $B'=\{a_i|i\in[k],\ \exists j\in[k]\ i\neq j,\ a_i=a_j\}$ and assume that

$$\mathbf{a}p_0(f(p_1, \dots, p_{i-1}, x_1, p_{i+1}, \dots, p_n)) \overset{*}{\underset{\mathbf{u}}{\Rightarrow}} r_0(\bar{f}(r_1, \dots, r_{i-1}, ux_1, r_{i+1}, \dots, r_n)),$$

$$\mathbf{a}p_0'(f(p_1', \dots, p_{i-1}', x_1, p_{i+1}', \dots, p_n')) \overset{*}{\underset{\mathbf{u}}{\Rightarrow}} r_0'(\bar{f}(r_1', \dots, r_{i-1}', ux_1, r_{i+1}', \dots, r_n')),$$

$$uq_0 \overset{*}{\underset{\mathbf{u}}{\Rightarrow}} q_0(\mathbf{v}(x_1, \dots, x_m)), \quad uq_0' \overset{*}{\underset{\mathbf{u}}{\Rightarrow}} q_0'(\mathbf{v}'(x_1, \dots, x_m)),$$

$$\mathbf{vq} \overset{*}{\underset{\mathbf{u}}{\Rightarrow}} \mathbf{s}, \quad \mathbf{v}'\mathbf{q}' \overset{*}{\underset{\mathbf{u}}{\Rightarrow}} \mathbf{s}',$$

$$\mathrm{rt}(\mathbf{s}) = \mathrm{rt}(\mathbf{s}')\in\bar{F}^m,$$

where $n>0$, $m\geqq 0$, $i\in[n]$, $p_0, p_0'\in\hat{T}_{F,1}$, $f\in F_n$, $\bar{f}=(f, C, \varphi, \psi)\in\bar{F}_n$, $p_j, p_j'\in T_F$, $r_0, r_0'\in\hat{T}_{F'',1}$, $r_j, r_j'\in T_{F''}$ $(j\in[n]-\{i\})$, $q_0, q_0'\in\hat{T}_{F,m}$, $\mathbf{q}, \mathbf{q}'\in T_F^m$, $\mathbf{s}, \mathbf{s}'\in T_{F''}^m$, $u\in U$, $\mathbf{v}, \mathbf{v}'\in U^m$.

If $|C|\geqq 2$ or $C\cap B'\neq\emptyset$ then $\tau_b(p_0(f(p_1, \dots, p_{i-1}, q_0(\mathbf{q}), p_{i+1}, \dots, p_n)))=\tau_b(p_0(f(p_1, \dots, p_{i-1}, q_0'(\mathbf{q}), p_{i+1}, \dots, p_n)))$ is valid for any $b\in B$. If $|C|=1$ and $C\cap B'=\emptyset$ then we have the same equality for any $b\in B-C$. Furthermore, $m\leqq|A|$.

Proof. Similar to the proof of Lemma 4.

By succesive applications of the previous two lemmata we obtain

Lemma 6. Assume that $\mathbf{ap} \overset{*}{\underset{\mathbf{A}}{\Rightarrow}} \mathbf{q}$ where $\mathbf{a}\in A^k$, $\mathbf{p}\in T_F^k$, $\mathbf{q}\in T_G^k$, $k\geqq 0$. If $\sigma_{\mathbf{a}}(p_1)\sim\dots\sim\sigma_{\mathbf{a}}(p_k)$ then there is a tree $p_0\in T_F$ with $\sigma_{\mathbf{a}}(p_0)\sim\sigma_{\mathbf{a}}(p_1)$ and $\mathbf{a}\mathbf{p}_0^k \overset{*}{\underset{\mathbf{A}}{\Rightarrow}} \mathbf{q}$. Furthermore, if $r\in\bigcap_{i=1}^{k}\mathrm{dom}\,\tau_{a_i}$ then $\overline{\mathrm{wd}}(\sigma_{\mathbf{a}}(r))\leqq|A|$.

Lemma 7. Let $\mathbf{a}\in A^k$ $(k\geqq 0)$, $f\in F_n$ $(n\geqq 0)$, $\mathbf{b}_{ij}\in A^{m_{ij}}$ $(m_{ij}\geqq 0,\ i\in[k],\ j\in[n])$ and $q_i\in\hat{T}_{G,m_i}$ $(i\in[k],\ m_i=\sum_{j=1}^{n} m_{ij})$. Assume that each of the productions $a_if\to q_i(\mathbf{b}_{i1}\mathbf{x}_1^{m_{i1}}, \dots, \mathbf{b}_{in}\mathbf{x}_n^{m_{in}})$ $(i\in[k])$ is in Σ. Furthermore, let $p_i, p_i'\in T_F$, $\mathbf{c}_i=(\mathbf{b}_{1i}, \dots, \mathbf{b}_{ki})$ $(i\in[n])$. Then $\sigma_{\mathbf{c}_i}(p_i)\sim\sigma_{\mathbf{c}_i}(p_i')$ $(i\in[n])$ implies $\sigma_{\mathbf{a}}(f(p_1, \dots, p_n))\sim\sigma_{\mathbf{a}}(f(p_1', \dots, p_n'))$.

Proof. The proof will be carried out in case of $n=1$ only. As $n=1$ we may simplify our notations: put $p=p_1$, $p'=p_1'$, $\mathbf{b}_i=\mathbf{b}_{i1}$ $(i\in[k])$, $\mathbf{c}=\mathbf{c}_1$. Moreover, let $B=\{a_i|i\in[k]\}$, $B'=\{a_i|i\in[k],\ \exists j\in[k]\ i\neq j,\ a_i=a_j\}$, $C=\{c_i|i\in[\sum_{j=1}^{k} m_j]\}$, $C'=\{c_i|i\in[\sum_{j=1}^{k} m_j],\ \exists i'\in[\sum_{j=1}^{k} m_j]\ i'\neq i,\ c_i=c_{i'}\}$.

As $p, p' \in \bigcap_{i=1}^{k} \bigcap_{j=1}^{m_i} \operatorname{dom} \tau_{b_{ij}}$ and the productions above exist also $f(p), f(p') \in \bigcap_{i=1}^{k} \operatorname{dom} \tau_{a_i}$. This implies that both $\sigma_a(f(p))$ and $\sigma_a(f(p'))$ are defined.

In the remaining part of the proof we shall make some transformations on the trees $f(\sigma_c(p))$ and $f(\sigma_c(p'))$ by the help of a deterministic top-down tree transducer $\mathbf{V} = (F'', V, F'', v_0, \Sigma_V)$. In this transducer $V = \{v_0\} \cup \{(D, \psi) | D \subsetneqq B, \psi: A \rightarrowtail A\}$ and Σ_V consisits of the following five types of rules:

(i) If $q_i = x_1$ for every $i \in [k]$ then

$$v_0 f \to f((\emptyset, \psi) x_1) \in \Sigma_V$$

where $\psi(a) = b$ if and only if $a = a_i$ and $b = b_{i1}$ for an index $i \in [k]$.

(ii) If $D = \{a_i | i \in [k], q_i \neq x_1\}$ is not empty then

$$v_0 f \to (f, D, \varphi, \psi)((D, \psi_1) x_1) \in \Sigma_V$$

where $\varphi: B \to P(A)$, $\varphi(a) = \{a\}$ $(a \in B)$; moreover, $\psi(a) = a$ if $a \in B$, $\psi(a)$ is undefined if $a \notin B$; $\psi_1(a) = a$ if $a \in C$, otherwise $\psi_1(a)$ is undefined.

(iii) $(D, \psi) g \to g((D, \psi) x_1, \ldots, (D, \psi) x_l) \in \Sigma_V$ for any $(D, \psi) \in V$ and $g \in F_l$ $(l \geqq 0)$.

(iv) If $(D, \psi) \in V$, $D' \subseteq C$ and either $|D| > 1$ or $D \cap B' \neq \emptyset$ or $\{a_i | \{b_{i1}, \ldots, b_{im_i}\} \cap D' \neq \emptyset\} \neq D$ then

$$(D, \psi)(g, D', \varphi', \psi') \to (g, D'', \varphi'', \psi'')((D'', \psi_1) x_1, \ldots, (D'', \psi_l) x_l) \in \Sigma_V$$

for any $(g, D', \varphi', \psi') \in \bar{F}_l$ $(l \geqq 0)$ with $\varphi': C \to P(A)$ where $D'' = \{a_i | i \in [k], \{b_{i1}, \ldots, b_{im_i}\} \cap D' \neq \emptyset\}$; $\varphi'': B \to P(A)$ and $\varphi''(a_i) = \bigcup_{j=1}^{m_i} \varphi'(b_{ij})$ $(i \in [k])$; $\psi'' = \psi \circ \psi'$ and $\psi_i(a) = b$ if and only if $a = b$ and a occurs in the right side of a rule $cg \to s \in \Sigma$ with $c \in \bigcup \varphi'(C)$.

(v) If $(D, \psi) \in V$, $D' \subseteq C$, furthermore $|D| = 1$, $D \cap B' = \emptyset$ and $\{a_i | i \in [k], \{b_{i1}, \ldots, b_{im_i}\} \cap D' \neq \emptyset\} = D$ then for every $(g, D', \varphi', \psi') \in \bar{F}_l$ with $\varphi': C \to P(A)$

$$(D, \psi)(g, D', \varphi', \psi') \to g((D, \psi_1) x_1, \ldots, (D, \psi_l) x_l) \in \Sigma_V$$

where $\psi_i = \psi \circ \eta_i$ and $\eta_i(a) = b$ if and only if $ag \to bx_i \in \Sigma$.

It can be seen that $\tau_V(f(\sigma_c(p))) = \sigma_a(f(p))$ and $\tau_V(f(\sigma_c(p'))) = \sigma_a(f(p'))$. On the other hand, by $\sigma_c(p) \sim \sigma_c(p')$ it follows that $\tau_V(f(\sigma_c(p))) \sim \tau_V(f(\sigma_c(p')))$. Therefore, $\sigma_a(f(p)) \sim \sigma_a(f(p'))$, as was to be proved.

We now turn to the definition of F'. For every integer $i \geqq 1$ let K_i denote the maximal number of occurences of the variable x_i in the right side of a rule in Σ. Put $K = \max\{1, K_i | i \in [v(F)]\}$, $F'_{nK} = F_n$ $(n \geqq 0)$ and $F'_m = \emptyset$ otherwise.

As it was mentioned we introduce two homomorphisms $\varrho \subseteq T_F \times T_{F'}$ and $\varrho' \subseteq T_{F'} \times T_F$ connecting T_F and $T_{F'}$. The rules defining ϱ are $f \to f(\mathbf{x}_1^K, \ldots, \mathbf{x}_n^K)$ $(f \in F_n,\ n \geqq 0)$, while the rules corresponding to ϱ' are $f \to f(x_{i_1}, \ldots, x_{i_n})$ $(f \in F_n,\ n \geqq 0)$ with $i_1 \in [K], \ldots, i_n \in [nK] - [(n-1)K]$. Observe that ϱ is deterministic and we have $\varrho'(\varrho(p)) = \{p\}$ for any $p \in T_F$.

We continue by defining the transducer $\mathbf{A}' = (F', A', G, a'_0, \Sigma')$. In this system $A'=\{(a, B, B')|a\in B, B\subseteq A, B'\subseteq B\}$, $a_0'=(a_0, \{a_0\}, \varnothing)$ and Σ' is the smallest set of rewriting rules with the following property.

Let $l>0$, $B=\{a_1, \ldots, a_l\}\subseteq A$, $B'=\{a_{m_1}, \ldots, a_{m_k}\}$ $(1\leqq m_1<\ldots<m_k\leqq l)$, $a=a_1$. Assume that the rules $a_i f\to q_i(\mathbf{a}_{i1}\mathbf{x}_1^{k_{i1}}, \ldots, \mathbf{a}_{in}\mathbf{x}_n^{k_{in}})$ are in Σ where $n\geqq 0$, $f\in F_n$, $k_{ij}\geqq 0$, $\mathbf{a}_{ij}\in A^{k_{ij}}$, $q_i\in\hat{T}_{G,k_{i1}+\ldots+k_{in}}$ $(i\in[l], j\in[n])$. Furthermore, let $r_j\in T(2|A|^2\|A\|^2, |A|)$, and set $R_j=\{p\in T_{F'}, |\varrho'(p)\leqq\sigma_{\mathbf{b}_j}^{-1}([r_j])\}$ $(j\in[n])$, where $\mathbf{b}_j=(\mathbf{a}_{1j}, \ldots, \mathbf{a}_{lj}, \mathbf{a}_{m_1 j}, \ldots, \mathbf{a}_{m_k j})$. R_j is regular by Lemma 2 and some results in [2]. Finally, denote by B_j the set of components of $\mathbf{b}_j$ and put $B_j'=\{b\in A|b$ occurs at least twice in $\mathbf{b}_j\}$ $(j\in[n])$, $c_{ij}=a_{1ij}$ $(i\in[n], j\in[k_{1i}])$, $k_i=k_{1i}$ $(i\in[n])$. Then the rule

$$((a, B, B')f\to q_1((c_{11}, B_1, B_1')x_1, \ldots, (c_{1k_1}, B_1, B_1')x_{k_1}, \ldots$$

$$\ldots, (c_{n1}, B_n, B_n')x_{(n-1)K+1}, \ldots, (c_{nk_n}, B_n, B_n')x_{(n-1)K+k_n}),$$

$$\underbrace{R_1, \ldots, R_1}_{K\text{-times}}, \ldots, \underbrace{R_n, \ldots, R_n}_{K\text{-times}})$$

is in Σ'.

Observe that with the definition above $\mathbf{A}'$ becomes a linear deterministic top-down tree transducer with regular look-ahead. Just as in case of $\mathbf{A}''$ we may treat any vector $\mathbf{a}\in A^l$ — but now with $l>0$ — as an element of A': if $\mathbf{a}\in A^l$ $(l>0)$ then identify $\mathbf{a}$ with (a_1, B, B') where $B=\{a_i|i\in[l]\}$, $B'=\{a_i|i\in[l], J_j\in[l]\ i\neq j, a_i=a_j\}$. Assume that $\mathbf{a}p\overset{*}{\underset{\mathbf{A}'}{\Rightarrow}}q$ $(p\in T_{F'}, q\in T_G)$. Then one can easily prove that $\varrho'(p)\subseteq$ $\subseteq\bigcap_{i=1}^{l}\operatorname{dom}\tau_{a_i}$. However, there is a much more close connection between $\mathbf{A}$ and $\mathbf{A}'$. This is shown by Lemmata 8 and 9. In these Lemmata we shall assume that $\mathbf{A}$ does not satisfy condition $(*)$.

Lemma 8. $\tau(\operatorname{dom}\tau)\subseteq\tau'(\operatorname{dom}\tau')$.

Proof. We shall prove that if $a_0 p_0\overset{*}{\underset{\mathbf{A}}{\Rightarrow}}q_0(\mathbf{a}\mathbf{x}_1^k)$ and $\mathbf{a}p^k\overset{*}{\underset{\mathbf{A}}{\Rightarrow}}\mathbf{q}$ where $k>0$, $p_0\in\hat{T}_{F,1}$, $p\in T_F$, $q_0\in\hat{T}_{G,k}$, $\mathbf{q}\in T_G^k$, $\mathbf{a}\in A^k$ then also $\mathbf{a}\varrho(p)\overset{*}{\underset{\mathbf{A}'}{\Rightarrow}}q_1$. From this the statement follows by taking $p_0=x_1$.

If $\operatorname{dp}(p)=0$, i.e. $p\in F_0$, then $\mathbf{a}\varrho(p)\overset{*}{\underset{\mathbf{A}'}{\Rightarrow}}q_1$ is obviously valid. We proceed by induction on $\operatorname{dp}(p)$. Therefore, suppose that $\operatorname{dp}(p)>0$ and the proof is done for trees with depth less than $\operatorname{dp}(p)$. Then $p=f(p_1, \ldots, p_n)$ where $n>0$, $f\in F_n$, $p_1, \ldots, p_n\in T_F$ and $\operatorname{dp}(p_i)<\operatorname{dp}(p)$ $(i\in[n])$. As the generalization to arbitrary n is straigthforward we shall deal with $n=1$ only. Since $\mathbf{a}p^k\overset{x}{\underset{\mathbf{A}}{\Rightarrow}}\mathbf{q}$ there exist rules $a_i f\to r_i(\mathbf{b}_i\mathbf{x}_1^{l_i})\in\Sigma$ $(i\in[k], l_i\geqq 0, r_i\in\hat{T}_{G,l_i}, \mathbf{b}_i\in A^{l_i})$ such that $\mathbf{b}_i\mathbf{p}_1^{l_i}\overset{*}{\underset{\mathbf{A}}{\Rightarrow}}\mathbf{s}_i$ and $q_i=r_i(\mathbf{s}_i)$ hold for some $\mathbf{s}_i\in T_G^{l_i}$. Put $l=l_1+\ldots+l_k$, $\mathbf{b}=(\mathbf{b}_1, \ldots, \mathbf{b}_k)$, $B=\{b|b$ occurs in $\mathbf{b}\}$, $B'=\{b|b$ occurs at least twice in $\mathbf{b}\}$. As $a_0 p_0(f(x_1))\overset{*}{\underset{\mathbf{A}}{\Rightarrow}}q_0(r_1(\mathbf{b}_1\mathbf{x}_1^{l_1}), \ldots, r_k(\mathbf{b}_k\mathbf{x}_1^{l_k}))$ and $\mathbf{b}\mathbf{p}_1^l\overset{*}{\underset{\mathbf{A}}{\Rightarrow}}(\mathbf{s}_1, \ldots, \mathbf{s}_k)$ we have that $\sigma_{\mathbf{b}}(p_1)$ is defined, $\sigma_{\mathbf{b}}(p_1)\in T(2|A|^2\|A\|^2, |A|)$ (cf. Lemmata 3 and 6). Set $R=\{p'\in T_{F'}|\varrho'(p')\subseteq\sigma_{\mathbf{b}}^{-1}([\sigma_{\mathbf{b}}(p_1)])$. By the construction of

$\mathbf{A}'$ we know that $(\mathbf{a}f \to r_1((b_{11}, B, B')x_1, \ldots, (b_{1l_1}, B, B')x_{l_1}), \underbrace{R, \ldots, R}_{K\text{-times}})$ is in Σ. Now, if $l_1 = 0$ then we get $\mathbf{a}\varrho(p) \overset{*}{\underset{\mathbf{A}}{\Rightarrow}} q_1$ immediately. If $l_1 > 0$ then we obtain $(b_{11}, B, B')\varrho(p_1) \overset{*}{\underset{\mathbf{A}'}{\Rightarrow}} s_{11}, \ldots, (b_{1l_1}, B, B')\varrho(p_1) \overset{*}{\underset{\mathbf{A}'}{\Rightarrow}} s_{1l_1}$ by the induction hypothesis. As $\varrho(p_1) \in R$ we again have $\mathbf{a}\varrho(p) \overset{*}{\underset{\mathbf{A}'}{\Rightarrow}} q_1$.

Lemma 9. $\tau'(\mathrm{dom}\,\tau') \subseteq \tau(\mathrm{dom}\,\tau)$.

Proof. We are going to show that if $\mathbf{a}p' \overset{*}{\underset{\mathbf{A}'}{\Rightarrow}} q$ where $\mathbf{a} \in A^l$ $(l > 0)$ $p' \in T_{F'}$, $q \in T_G$ then there exist trees $r \in T_{F''}$ and $p \in \sigma_{\mathbf{a}}^{-1}([r])$ with $\varrho'(p') \subseteq \sigma_{\mathbf{a}}^{-1}([r])$ and $a_1 p \overset{*}{\underset{\mathbf{A}}{\Rightarrow}} q$. If $\mathrm{dp}(p') = 0$ then it is trivial: take $p = p'$, $r = \sigma_{\mathbf{a}}(p)$. Assume now that this statement is valid for trees with depth less than $\mathrm{dp}(p')$ and $\mathrm{dp}(p') \geqq 1$. Then $p' = f(p'_1, \ldots, p'_{nK})$ $(n > 0)$ with $\mathrm{dp}(p'_1), \ldots, \mathrm{dp}(p'_{nK}) < \mathrm{dp}(p')$. We shall restrict ourselves to the case $n = 1$. Since $\mathbf{a}p' \overset{*}{\underset{\mathbf{A}'}{\Rightarrow}} q$ we get

$$(\mathbf{a}f \to q_0((b_1, B, B')x_1, \ldots, (b_k, B, B')x_k), \underbrace{R, \ldots, R}_{K\text{-times}}) \in \Sigma',$$

$$(b_i, B, B')p'_i \overset{*}{\underset{\mathbf{A}'}{\Rightarrow}} q_i \quad (i \in [k]), \quad p'_i \in R \quad (i \in [K]),$$

$$q = q_0(q_1, \ldots, q_k)$$

for some k $(0 \leqq k \leqq K)$, $b_1, \ldots, b_k \in A$, $B, B' \subseteq A$ with $\{b_1, \ldots, b_k\} \subseteq B$, $B' \subseteq B$, $q_0 \in \hat{T}_{G,k}$, $q_1, \ldots, q_k \in T_G$ and a regular forest $R = \{s \in T_{F'} \mid \varrho'(s) \subseteq \sigma_{\mathbf{c}}^{-1}([r_1])\}$ where $r_1 \in T_{F''}$ and $\mathbf{c}$ is an arbitrary vector containing one component c_i for each element c_i of B and a distinct component c_j for each element c_j of B'. We have by the definition of $\mathbf{A}'$ that $a_1 f \to q_0(b_1 x_1, \ldots, b_k x_1) \in \Sigma$. Furthermore, as $\varrho'(p'_1), \ldots, \varrho'(p'_K) \subseteq \sigma_{\mathbf{c}}^{-1}([r_1])$, by Lemma 7 we have $\varrho'(f(p'_1, \ldots, p'_K)) \subseteq \sigma_{\mathbf{a}}^{-1}([r])$ for a suitable $r \in T_{F''}$.

If $k = 0$ then let $\bar{p} \in \varrho'(p'_1)$ be arbitrary, $p = f(\bar{p})$. $a_1 p \overset{*}{\underset{\mathbf{A}}{\Rightarrow}} q$ follows obviously. By $\bar{p} \in \varrho'(p'_1)$ also $f(\bar{p}) \in \varrho'(f(p'_1, \ldots, p'_K))$. Thus, $p = f(\bar{p}) \in \sigma_{\mathbf{a}}^{-1}([r])$.

If $k > 0$ then there are trees $p_1, \ldots, p_k \in \sigma_{\mathbf{c}}^{-1}([r_1])$ with $b_1 p_1 \overset{*}{\underset{\mathbf{A}}{\Rightarrow}} q_1, \ldots, b_k p_k \overset{*}{\underset{\mathbf{A}}{\Rightarrow}} q_k$. From this, by an application of Lemma 6, it follows that there is a tree $\bar{p} \in \sigma_{\mathbf{c}}^{-1}([r_1])$ with $b_1 \bar{p} \overset{*}{\underset{\mathbf{A}}{\Rightarrow}} q_1, \ldots, b_k \bar{p} \overset{*}{\underset{\mathbf{A}}{\Rightarrow}} q_k$. Put $p = f(\bar{p})$. Again, we have $a_1 p \overset{*}{\underset{\mathbf{A}}{\Rightarrow}} q$. On the other hand, $p \in \sigma_{\mathbf{a}}^{-1}([r])$. Indeed, let $\bar{p}_1 \in \varrho'(p'_1)$ be arbitrary. Then, as $\sigma_{\mathbf{c}}(\bar{p}) \sim \sigma_{\mathbf{c}}(\bar{p}_1)$, $\sigma_{\mathbf{a}}(f(\bar{p})) \sim \sigma_{\mathbf{a}}(f(p_1))$ follows by Lemma 7. By $f(\bar{p}_1) \in \sigma_{\mathbf{a}}^{-1}([r])$ this means that $f(\bar{p}) \in \sigma_{\mathbf{a}}^{-1}([r])$.

Now we are ready to state the main result of this section:

Theorem 1. A deterministic top-down tree transducer $\mathbf{A}$ preserves regularity if and only if $(*)$ is not satisfied by $\mathbf{A}$. The regularity preserving property of deterministic top-down transducers is decidable.

Proof. The necessity of the first statement of our Theorem is valid by Lemma 1. To prove the converse suppose that $\mathbf{A}=(F, A, G, a_0, \Sigma)$ does not satisfy condition $(*)$, and take a regular forest $R \subseteq T_F$. R is recognizable by a deterministic tree automaton $\mathbf{B}=(F, B, B_0)$. Without loss of generality we may assume that $\mathbf{B}$ is connected, i.e., for any state $b \in B$ there is a tree $p \in T_F$ with $(p)_{\mathbf{B}}=b$.

First let B_0 be a singleton set, say $B_0=\{b_0\}$, and take the deterministic top-down tree transducer $\mathbf{A}'=(H, A\times B, G, (a_0, b_0), \Sigma')$ where $H_n=\{(f, b_1, \ldots, b_n) \mid f\in F_n, b_1, \ldots, b_n \in B\}$ $(n \geqq 0)$

$$\Sigma' = \{(a, b)(f, b_1, \ldots, b_n) \rightarrow q((a_1, b_{i_1})x_{i_1}, \ldots, (a_m, b_{i_m})x_{i_m}) \mid$$

$$|m, n \geqq 0, \quad a, a_1, \ldots, a_m \in A, \quad b_1, \ldots, b_n \in B, \quad i_1, \ldots, i_m \in [n],$$

$$af \rightarrow q(a_1 x_{i_1}, \ldots, a_m x_{i_m}) \in \Sigma, \quad b = (f)_{\mathbf{B}}(b_1, \ldots, b_n)\}.$$

It is not difficult to see that $\tau_{\mathbf{A}}(R)=\tau_{\mathbf{A}'}(\mathrm{dom}\, \tau_{\mathbf{A}'})$. On the other hand $\mathbf{A}'$ does not satisfy $(*)$. By Lemmata 8 and 9, and the fact that linear top-down transducers with regular look-ahead preserve regularity (cf. [2], [3]), this implies that $\tau_{\mathbf{A}}(R)$ is regular.

The general case, i.e. when B_0 is arbitrary, is reducible to the previous one. Indeed, if $B=\{b_1, \ldots, b_n\}$ then put $\mathbf{B}_i=(F, B, \{b_i\})$, $R_i=T(\mathbf{B}_i)$ $(i\in[n])$. Obviously, $\tau_{\mathbf{A}}(R)=\bigcup_{i=1}^{n} \tau_{\mathbf{A}}(R_i)$. As all the $\tau_{\mathbf{A}}(R_i)$ are regular and regular forests are closed under union, it follows that $\tau_{\mathbf{A}}(R)$ is regular, as well.

The second statement of Theorem 1 is a consequence of the first one because it is decidable whether $(*)$ is satisfied by $\mathbf{A}$.

As every uniform deterministic top-down transducer is equivalent to a nondeterministic bottom-up transducer, by the characterization theorem for regularity preserving bottom-up transducers in [4], it follows that a uniform deterministic top-down transducer preserves regularity if and only if it is equivalent to a linear bottom-up transducer. In general, we do not know any similar characterization for regularity preserving deterministic top-down transducers.

3. Nondeterministic top-down tree transducers

In this section we prove

Theorem 2. The regularity preserving property of nondeterministic top-down tree transducers is undecidable.

Proof. Let H be an arbitrary type containing unary operational symbols only. Take a Post Correspondence Problem $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ $(\boldsymbol{\alpha}, \boldsymbol{\beta} \in H^{+m}, m>0)$ and choose l in such a way that $|\alpha_i|, |\beta_i|<l$ $(i\in[m])$. Set $F_0=\{\#\}$, $F_1=[m]$ $([m]\cap H=\emptyset)$, $F=F_0\cup F_1$, $G_0=F_0$, $G_1=F_1\cup H\cup\{f\}$ $(f\notin F_1\cup H)$, $G_2=\{g\}$, $G=G_0\cup G_1\cup G_2$. We shall give a top-down tree transformation $\tau\subseteq T_F\times T_G$ such that τ preserves regularity if and only if $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ has no solution.

Consider the top-down transducer $\mathbf{A}_1=(F, \{a_0, a_1, a_2, b_1, b_2, b_3\}, G, a_0, \Sigma)$ with Σ consisiting of the rules from (1) to (8) where $i\in[m]$:

(1) $a_0 i \to a_0 x_1$,

(2) $a_0 i \to g(f(a_1 x_1), \alpha_i(b_1 x_1))$, [2]

$a_0 i \to g(f(a_1 x_1), w(b_2 x_1)) \quad (w\in H^*, |w| \leqq |\alpha_i|, w \neq \alpha_i)$,

(3) $a_1 i \to f(a_1 x_1), \quad a_1 \# \to \#$,

(4) $b_1 i \to \alpha_i(b_1 x_1), \quad b_1 i \to w(b_2 x_1) \quad (w\in H^*, |w| \leqq \alpha_i, w \neq \alpha_i)$,

(5) $b_2 i \to w(b_2 x_1) \quad (w\in H^*, |w| \leqq \alpha_i, w \neq \alpha_i), \quad b_2 \# \to \#$,

(6) $a_0 i \to g(a_2 x_1, w(b_3 x_1)) \quad (w\in H^*, 1 \leqq |w| \leqq l)$,

$a_0 i \to g(f(a_2 x_1), w(b_3 x_1)) \quad (w\in H^*, |\alpha_i| < |w| \leqq l)$,

(7) $a_2 i \to a_2 x_1, \quad a_2 i \to f(a_2 x_1), \quad a_2 \# \to \#$,

(8) $b_3 i \to w(b_3 x_1) \quad (w\in H^*, |\alpha_i| \leqq |w| \leqq l), b_3 \# \to \#$.

Denote $\tau_{\mathbf{A}_1}$ by τ_1. It can be seen that τ_1 consists of all pairs $(i_1 \dots i_k(\#), g(f^{k-j}(\#), w(\#)))$ where $k\geqq 1$, $0\leqq j\leqq k$, $w\in H^*$, $0\leqq |w|\leqq kl$ and $w\neq\alpha_{i_{j+1}} \dots \alpha_{i_k}$. Similarly, a top-down tree transducer $\mathbf{A}_2$ inducing τ_2 can be constructed with τ_2 containing the same pairs as τ_1 with the exception that $w\neq\beta_{i_{j+1}}\dots\beta_{i_k}$. Taking the disjoint sum of $\mathbf{A}_1$ and $\mathbf{A}_2$ we obtain a top-down transducer $\mathbf{A}$ inducing $\tau=\tau_1\cup\tau_2$.

Assume that $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ has a solution. Then let $i_1\dots i_k$ be a solution to $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ with minimal length. Put $L=\{(i_1\dots i_k)^n(\#)|n\geqq 0\}$, $w=\alpha_{i_1}\dots\alpha_{i_k} (=\beta_{i_1}\dots\beta_{i_k})$, $T=\overline{\tau(L)}\cap \cap\{g(f^r(\#), v(\#))|r\geqq 0, v\in H^*\}$, $R=\{g(f^{kn}(\#), w^n(\#))|n\geqq 0\}$. We are going to show that $T=R$. As the class of regular forests is closed under complementation and meet, furthermore, the forest $\{g(f^r(\#), v(\#))|r\geqq 0, v\in H^*\}$ is regular while R is not, from this follows that $\tau(L)$ is not regular. Since L is regular this implies that τ does not preserve regularity.

Suppose that $g(f^{kn}(\#), w^n(\#))\in\tau(L)$. Then there exists an integer r $(0\leqq n\leqq r)$ with $g(f^{kn}(\#), w^n(\#))\in\tau((i_1\dots i_k)^n(\#))$. Therefore, either $w^n \neq (\alpha_{i_1}\dots\alpha_{i_k})^n$ or $w^n\neq(\beta_{i_1}\dots\beta_{i_k})^n$. As $i_1\dots i_k$ is a solution to $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ both cases yield a contradiction. Thus, $R\subseteq T$. To prove the converse suppose that $g(f^r(\#), v(\#))\notin\{g(f^{kn}(\#), w^n(\#))|n\geqq 0\}$ $(r\geqq 0, v\in H^*)$. Let $n\geqq\max\{r, |v|/l\}$ be the least integer divisible by k, $j_1\dots j_n=(i_1\dots i_k)^{n/k}$. If r is a multiple of k, say $r=kt$, then $v\neq w^t$, i.e. $v\neq\alpha_{j_{r+1}}\dots\alpha_{j_n}$. If r is not a multiple of k then, as $i_1\dots i_k$ was a minimal solution to $(\boldsymbol{\alpha}, \boldsymbol{\beta})$, $j_{r+1}\dots j_n$ is not a solution to $(\boldsymbol{\alpha}, \boldsymbol{\beta})$. Therefore, either $v\neq\alpha_{j_{r+1}}\dots\alpha_{j_n}$ or

[2] If F is a unary type and $v=f_1\dots f_k\in F^*$ then we denote by v the tree $f_1(\dots(f_k(x_1))\dots)\in T_{F,1}$ as well.

$v \neq \beta_{j_{r+1}} \dots \beta_{j_n}$. Moreover, as $n \geqq |v|/l$, in both cases $|v| \leqq ln$. This together with $n>0$ means that $g(f^r(\#), v(\#)) \in \tau(j_1 \dots j_n(\#)) \subseteq \tau(L)$, as was to be proved.

Next assume that $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ has no solution. Then $\tau(L)=\{g(f^r(\#), v(\#)) | r \geqq 0, v \in H^*\} - \{g(\#, \#)\}$ holds for any infinite $L \subseteq T_F$. Consequently, **A** preserves regularity.

DEPT. OF COMPUTER SCIENCE
A. JÓZSEF UNIVERSITY
ARADI VÉRTANÚK TERE 1
SZEGED, HUNGARY
H-6720

References

[1] ÉSIK, Z., Decidability results concerning tree transducers I., *Acta Cybernet.*, v. 5, 1980, pp. 1—20.
[2] ENGELFRIET, J., Bottom-up and top-down tree transformations, A comparison, *Math. Systems Theory*, v. 9, 1975, pp. 198—231.
[3] ENGELFRIET, J., Top-down tree transducers with regular lookahead, *Math. Systems Theory*, v. 10, 1977, pp. 289—303.
[4] GÉCSEG, F., Tree transformations preserving recognizability, Proc., Finite algebra and multiple-valued logic, North Holland, 1981, pp. 251—273.

(Received Feb. 7, 1983)