An infinite hierarchy of tree transformations in the class $\mathcal{ND}\mathcal{R}$

S. VÁGVÖLGYI and Z. FÜLÖP

Introduction

Let $S=\{\mathcal{DR}, \mathcal{LDR}, \mathcal{NDR}, \mathcal{LNDR}, \mathcal{H}, \mathcal{LH}, \mathcal{NH}\}$ where $\mathcal{DR}$ is the class of all deterministic root-to-frontier tree transformations, $\mathcal{H}$ is the class of all homomorphism tree transformations, moreover, for both $\mathcal{DR}$ and $\mathcal{H}$, their linear, nondeleting and linear-nondeleting subclasses are denoted by prefixing them by $\mathcal{L}$, $\mathcal{N}$ and $\mathcal{LN}$, respectively. Let $[S]$ be the set of classes of tree transformations generated by S with composition $\circ$: $[S]=\{\mathcal{K}_1\circ\ldots\circ\mathcal{K}_n | n\geqq 1,\ \mathcal{K}_i\in S,\ 1\leqq i\leqq n\}$. The set $[S]$ was introduced and examined in [1] where several equalities and inclusions were obtained with respect to elements of $[S]$. However, the question that whether $[S]$ is a finite or an infinite set was only raised and not answered.

In Section 2 of this paper we show that, in fact, $[S]$ is infinite by proving that $(\mathcal{LNDR}\circ\mathcal{NH})^m\subset(\mathcal{LNDR}\circ\mathcal{NH})^{m+1}$ for each $m\geqq 1$. This infinite proper hierarchy was already suggested by Theorem 12 of [1].

It is well known that $\mathcal{NDR}$ is closed under composition (proof, for example, in [1]). Thus we have $(\mathcal{LNDR}\circ\mathcal{NH})^m\subset\mathcal{NDR}$ for each $m\geqq 1$. In the second half of Section 2 we show that the stronger proper inclusion $\bigcup_{m=1}^{\infty}(\mathcal{LNDR}\circ\mathcal{NH})^m\subset$ $\subset\mathcal{NDR}$ is also valid.

The paper, apart from some simple reference to [1], is self-containing. Both in [1] and this paper, most of the notions and notations are adopted from [2].

1. Notions and notations

For an arbitrary set Y, we denote by Y^* the free monoid generated by Y, with empty word λ. The prefix ordering $\leqq$ in Y^* is meant as usual: for any $\alpha, \beta\in Y^*$, $\alpha\leqq\beta$ if and only if α is a prefix of β, that is, there exists a $\gamma\in Y^*$ such that $\beta=\alpha\gamma$. The relation $\alpha<\beta$ is defined by $\alpha\leqq\beta$ and $\alpha\neq\beta$.

The set of nonnegative integers is denoted by N. For each $n \in N$, $[n]$ denotes the set $\{1, ..., n\}$. Thus $[0]=\emptyset$.

By a ranked alphabet we mean an ordered pair (F, ν) where F is a finite set and $\nu\colon F \to N$ is the arity function. Elements of F are called function symbols, more exactly, if $f \in F$ and $\nu(f)=n$ then f is an n-ary function symbol. For any $n \in N$ we put $F_n=\{f \in F \mid \nu(f)=n\}$. Hence, for any ranked alphabet (F, ν), we have the equivalent notation $F=\bigcup_{n \in N} F_n$, where F_n are pairwise disjoint finite sets.

Let $F=\bigcup_{n \in N} F_n$ be a ranked alphabet and Y be a set, disjoint with F. Then the set of all terms or trees over Y of type F is defined as the smallest set $T_F(Y)$ satisfying:

(a) $Y \subseteq T_F(Y)$ and

(b) $f(p_1, ..., p_n) \in T_F(Y)$ whenever $f \in F_n$ and $p_1, ..., p_n \in T_F(Y)$.

For $f(\ )$ we write f. If $Y=\emptyset$ then $T_F(Y)$ is written as T_F.

We shall need a few of the usual functions on the elements of $T_F(Y)$: for any $p \in T_F(Y)$ the frontier $\mathrm{fr}(p) \in Y^*$, the set of subtrees or subterms $\mathrm{sub}(p) \subseteq T_F(Y)$, the paths $\mathrm{path}(p) \subseteq N^*$ and for each $m \in N$ the m-rank $\mathrm{rn}_m(p) \in N$ of p are defined by induction as follows:

(a) if $p \in Y$ then

$$\mathrm{fr}(p) = p, \quad \mathrm{sub}(p) = \{p\}, \quad \mathrm{path}(p) = \{\lambda\} \quad \text{and} \quad \mathrm{rn}_m(p) = 0;$$

(b) if $p = f(p_1, ..., p_n)$ for some $n \in N$, $f \in F_n$ and $p_1, ..., p_n \in T_F(Y)$ then

$$\mathrm{fr}(p) = \mathrm{fr}(p_1) \dots \mathrm{fr}(p_n),$$

$$\mathrm{sub}(p) = \Big(\bigcup_{i \in [n]} \mathrm{sub}(p_i)\Big) \cup \{p\},$$

$$\mathrm{path}(p) = \{\lambda\} \cup \{i\alpha \mid i \in [n],\ \alpha \in \mathrm{path}(p_i)\} \quad \text{and}$$

$$\mathrm{rn}_m(p) = \begin{cases} \sum_{i \in [n]} \mathrm{rn}_m(p_i) & \text{if } n \neq m \\ 1 + \sum_{i \in [n]} \mathrm{rn}_m(p_i) & \text{if } n = m. \end{cases}$$

We mention that $\mathrm{rn}_m(p)$ means the number of occurrences of the m-ary function symbols in p. Moreover we define $\mathrm{rn}(p) = \sum_{m \in N} \mathrm{rn}_m(p)$.

Now let $p \in T_F(Y)$ and $\alpha \in \mathrm{path}(p)$. We introduce the notion of the subtree $\mathrm{str}(p, \alpha)$ and the symbol $\mathrm{lab}(p, \alpha)$ of p determined by α, moreover, the two length $|\alpha|_2$ of α in p in the following way:

(a) if $p \in Y$ then
$\mathrm{str}(p, \alpha) = p$, $\mathrm{lab}(p, \alpha) = p$ and $|\alpha|_2 = 0$;

(b) if $p=f(p_1, ..., p_n)$ for some $n \in N$, $f \in F_n$ and $p_1, ..., p_n \in T_F(Y)$ then α is either λ or of the form $i\alpha'$ for some $i \in [n]$ and $\alpha' \in \mathrm{path}(p_i)$. Thus we define

$$\operatorname{str}(p, \alpha) = \begin{cases} p & \text{if } \alpha = \lambda \\ \operatorname{str}(p_i, \alpha') & \text{if } \alpha = i\alpha', \end{cases}$$

$$\operatorname{lab}(p, \alpha) = \begin{cases} f & \text{if } \alpha = \lambda \\ \operatorname{lab}(p_i, \alpha') & \text{if } \alpha = i\alpha', \end{cases}$$

$$|\alpha|_2 = \begin{cases} 0 & \text{if } \alpha = \lambda \text{ and } n < 2, \\ 1 & \text{if } \alpha = \lambda \text{ and } n \geqq 2, \\ |\alpha'|_2 & \text{if } \alpha = i\alpha' \text{ and } n < 2, \\ 1+|\alpha'|_2 & \text{if } \alpha = i\alpha' \text{ and } n \geqq 2. \end{cases}$$

We note that in this latter definition $|\alpha'|_2$ is meant in p_i. We mention what the above three functions informally mean. It is well known that p can be considered as an ordered tree labelled by elements of $F \cup Y$, moreover α can be thought of as a path leading from the root to a node x of p. Now, $\operatorname{str}(p, \alpha)$ is the subtree of p the root of which is x, $\operatorname{lab}(p, \alpha)$ is the symbol in $F \cup Y$ x is labelled by, finally $|\alpha|_2$ is the number of the occurrences of function symbols with arity $m \geqq 2$ along the path α. We also note that α may be in $\operatorname{path}(q)$ for some $q \neq p$ and $|\alpha|_2$ in p may differ from $|\alpha_2|$ in q. However it will always be clear from the context in what p $|\alpha|_2$ is meant.

The countably infinite set $X = \{x_1, x_2, \ldots\}$ of variable symbols will be kept fix throughout this paper. The set of the first m elements $x_1, \ldots, x_m$ of X is denoted by X_m. The set $T_F(X_m)$ will be written as $T_{F,m}$.

$\hat{T}_{F,m}$ is the linear-nondeleting subset of $T_{F,m}$: for $p \in T_{F,m}$, $p \in \hat{T}_{F,m}$ iff each x_i appears exactly once in p $(i \in [m])$.

For $p, q \in T_{F,m}$ and $i \in [m]$, by the i product $p \cdot_i q$ of p by q we mean the tree obtained from p by substituting each occurrence of x_i in p by q.

Let $p \in T_{F,m}$ and $y_1, \ldots, y_m \in Y$. We denote by $p(y_1, \ldots, y_m)$ the tree obtained from p by substituting each occurrence of x_i in p by y_i for each $i \in [m]$. Of course we have $p(y_1, \ldots, y_m) \in T_F(Y)$.

We introduce one more definition concerning $T_{F,m}$. For $p \in T_{F,m}$ and $i \in [m]$, the set of i paths $\operatorname{path}_i(p)$ of p is given as follows:

(a) if $p = x_j$ for some $j \in [m]$ then

$$\operatorname{path}_i(p) = \begin{cases} \{\lambda\} & \text{if } i = j \\ \emptyset & \text{if } i \neq j, \end{cases}$$

(b) if $p = f(p_1, \ldots, p_n)$ for some $n \geqq 0$, $f \in F_n$ and $p_1, \ldots, p_n \in T_{F,m}$ then

$$\operatorname{path}_i(p) = \{j\alpha \mid j \in [n],\ \alpha \in \operatorname{path}_i(p_j)\}.$$

It is clear that $\operatorname{path}_i(p) \subseteq \operatorname{path}(p)$, moreover $\operatorname{path}_i(p)$ consists of all the elements of $\operatorname{path}(p)$ leading from the root to a terminal node of p labelled by x_i.

A tree transformation τ is defined as a subset of $T_F \times T_G$ where F and G are arbitrary ranked alphabets. In this way, τ can alternatively be considered as a relation from T_F to T_G.

For the sake of convenient proofs, we introduce the concept of the extended tree transformation. It is a subset τ of $T_F(X) \times T_G(X)$.

Since (extended) tree transformations are in fact relations, for any (extended) tree transformations τ and σ, the domain $\operatorname{dom}\tau$ and the composition $\tau\circ\sigma$ of τ and σ are defined as it is usual for relations. Moreover, for any two classes $\mathcal{K}_1$ and $\mathcal{K}_2$ of tree transformations we put:

$$\mathcal{K}_1\circ\mathcal{K}_2 = \{\tau_1\circ\tau_2 | \tau_1\in\mathcal{K}_1 \text{ and } \tau_2\in\mathcal{K}_2\} \text{ and}$$

$$\mathcal{K}_1^n = \begin{cases}\mathcal{K}_1 & \text{if } n=1\\ \mathcal{K}_1^{n-1}\circ\mathcal{K}_1 & \text{if } n>1.\end{cases}$$

We are interested only in tree transformations which can be induced by deterministic root-to-frontier tree transducers.

A deterministic root-to-frontier tree transducer (DR transducer in the sequel) is a system

$$\mathfrak{A} = (F, A, G, P, a_0) \text{ where} \tag{1}$$

(a) F and G are ranked alphabets;
(b) A, the state set of $\mathfrak{A}$, is a ranked alphabet consisting of 1-ary function symbols, disjoint with F, G and X;
(c) a_0, the initial state of $\mathfrak{A}$, is a distinguished element of A;
(d) P is a finite set of so called rewriting rules (or simply rules) of the form

$$af(x_1, \ldots, x_n) \to q \tag{2}$$

where $a\in A$, $n\geqq 0$, $f\in F_n$ and $q\in T_G(AX_n)$;
(e) different rules of P have different left-hand sides.

We mention that above and in what follows we use the following notations. If A is the state set of a DR transducer and T is a set of terms then $AT=\{a(t)|a\in A,\ t\in T\}$. Moreover, for any $a\in A$ and $t\in T$, $a(t)$ is written as at.

Then it is clear that each rule (2) of P can also be written in both of the following two forms:

$$af(x_1, \ldots, x_n) \to \bar{q}(a_1x_{i_1}, \ldots, a_mx_{i_m}) \tag{3}$$

for some $m\geqq 0$, $\bar{q}\in\hat{T}_{G,m}$, $a_j\in A$ and $x_{i_j}\in X_n$ $(j\in[m])$;

$$af(x_1, \ldots, x_n) \to \bar{q}(a_{1_1}x_1, \ldots, a_{1_{m_1}}x_1, \ldots, a_{n_1}x_n, \ldots, a_{n_{m_n}}x_n) \tag{4}$$

where $m_i\geqq 0$, $a_{i_j}\in A$ $(i\in[n], j\in[m_i])$ and $\bar{q}\in\hat{T}_{G,m}$ $(m=m_1+\ldots+m_n)$.

Next we show how $\mathfrak{A}$ can be used to rewrite (or transform) terms of T_F to terms of T_G. To this end, we define the relation $\underset{\mathfrak{A}}{\Rightarrow}$ called direct derivation on the set $T_G(AT_F(X))$ in the following manner: for $p, q\in T_G(AT_F(X))$, $p\underset{\mathfrak{A}}{\Rightarrow} q$ if and only if q can be obtained from p by replacing an occurrence of a subtree of the form $af(p_1, \ldots, p_n)$ in p by the tree $\bar{q}(a_1p_{i_1}, \ldots, a_mp_{i_m})$ and the rule (3) is in P. The reflexve-transitive closure of $\underset{\mathfrak{A}}{\Rightarrow}$ is denoted by $\overset{*}{\underset{\mathfrak{A}}{\Rightarrow}}$ and called derivation. The tree transformation induced by $\mathfrak{A}$ with the state $a\in A$ is introduced as

$$\tau_{\mathfrak{A}(a)} = \{(p, q)|p\in T_F, q\in T_G \text{ and } ap\overset{*}{\underset{\mathfrak{A}}{\Rightarrow}} q\},$$

moreover, the tree transformation $\tau_{\mathfrak{A}}$ induced by $\mathfrak{A}$ is meant $\tau_{\mathfrak{A}(a_0)}$:

$$\tau_{\mathfrak{A}} = \{(p, q)|p\in T_F, q\in T_G \text{ and } a_0p\overset{*}{\underset{\mathfrak{A}}{\Rightarrow}} q\}.$$

Now we define the extended tree transformation induced by $\mathfrak{A}$. To this end, we need the following concept. Let $q' \in T_G(AX)$. We say that $q \in T_G(X)$ belongs to q' if it satisfies the following conditions:

(a) if $q' = ax_i$ for some $a \in A$ and $i \in N$ then $q = x_i$,

(b) if $q' = g(q'_1, ..., q'_n)$ for some $n \geqq 0$, $g \in G_n$ and $q'_1, ..., q'_n \in T_G(AX)$ then $q = g(q_1, ..., q_n)$ where q_j belongs to q'_j for each $j \in [n]$.

Informally, we say that q belongs to q' if and only if q can be obtained from q' by substituting each subtree of the form ax_i of q' by x_i.

The extended tree transformation $\tau'_{\mathfrak{A}(a)}$ induced by $\mathfrak{A}$ with the state $a \in A$ is given as follows: for any $p \in T_F(X)$ and $q \in T_G(X)$, $(p, q) \in \tau'_{\mathfrak{A}(a)}$ if and only if $\exists q' \in T_G(AX)$ such that $ap \overset{*}{\underset{\mathfrak{A}}{\Rightarrow}} q'$ and q belongs to q'. The extended tree transformation $\tau'_{\mathfrak{A}}$ induced by $\mathfrak{A}$ is defined as $\tau'_{\mathfrak{A}(a_0)}$.

We say that a tree transformation τ can be induced by some DR transducer $\mathfrak{A}$ if $\tau = \tau_{\mathfrak{A}}$ holds.

The tree transformations which can be induced by DR transducers are in fact partial mappings. This follows from (e) of the definition of the DR transducer.

Next we introduce some restrictions on DR transducers. We say that a DR transducer (1) is

(a) totally defined if for each $a \in A$ and $f \in F$ there is a rule (2) in P;

(b) linear (L) if for each rule (4) of P and $i \in [n]$, $m_i \leqq 1$;

(c) nondeleting (N) if for each rule (4) of P and $i \in [n]$, $m_i \geqq 1$;

(d) linear-nondeleting (LN) if it is linear and nondeleting;

(e) uniform (U) if for each rule (4) of P and $i \in [n]$, $a_{i_1} = ... = a_{i_{m_i}}$.

It is obvious that if a DR transducer (1) is a UDR transducer then each rule of P can also be written in the form $af(x_1, ..., x_n) \to \bar{q}(a_1 x_1, ..., a_n x_n)$ for some $\bar{q} \in T_{G,n}$ and $a_1, ..., a_n \in A$. Any LDR transducer is a UDR transducer, too.

A DR transducer (1) is an H transducer if it is totally defined and has only one state, i.e., $A = \{a_0\}$ holds. The LH, NH and LNH subclasses of the class of H transducers are defined in a natural way. Each H transducer is a UDR transducer, by definition.

The class of all tree transformations which can be induced by K transducers is denoted by $\mathcal{K}$ where K stands for any of DR, LDR, NDR, LNDR, UDR, H, LH, NH and LNH.

2. The problems and the solutions

Following [1], let S consist of $\mathcal{DR}$, $\mathcal{H}$ and their linear, nondeleting and linear nondeleting subclasses, that is, let $S = \{\mathcal{DR}, \mathcal{LDR}, \mathcal{NDR}, \mathcal{LNDR}, \mathcal{H}, \mathcal{LH}, \mathcal{NH}\}$. Moreover, define $[S]$ as the set of all the classes of tree transformations which can be obtained as compositions of elements of S: $[S] = \{\mathcal{K}_1 \circ ... \circ \mathcal{K}_n | n \geqq 1, \mathcal{K}_i \in S, 1 \leqq i \leqq n\}$.

In [1], it was raised the problem that whether $[S]$ is an infinite set. We shall prove that $[S]$ is infinite by showing that $[S]$ contains an infinite proper hierarchy of classes of tree transformations. Namely, we prove, in Theorem 3, that $(\mathcal{LNDR} \circ \mathcal{NH})^m \subset \subset (\mathcal{LNDR} \circ \mathcal{NH})^{m+1}$ for each $m \geqq 1$.

In connection with this hierarchy one more problem can be raised. By definition, $\mathcal{LNDR}$ and $\mathcal{NH}$ are subclasses of $\mathcal{NDR}$, moreover it is not difficult to see that $\mathcal{NDR}$ is closed under composition (a proof is given, e.g., in [1]). These, together with the infinite proper hierarchy mentioned above yield the proper inclusion $(\mathcal{LNDR} \circ \mathcal{NH})^m \subset \mathcal{NDR}$ for each $m \geqq 1$. In the second half of this section we show that the proper inclusion $\bigcup_{m=1}^{\infty} (\mathcal{LNDR} \circ \mathcal{NH})^m \subset \mathcal{NDR}$ also holds. Namely, in Lemma 17, we give an NDR transducer $\mathfrak{A}$ for which there does not exist m with $\tau_{\mathfrak{A}} \in (\mathcal{LNDR} \circ \mathcal{NH})^m$.

We set out to solve the first problem.

First we make a trivial observation on UNDR transducers. Let $\mathfrak{A} = (F, A, G, P, a_0)$ be a UNDR transducer and the rule $af(x_1, \ldots, x_n) \to q(a_1 x_1, \ldots, a_n x_n)$ in P. Then for each $j \in [n]$ and $\gamma \in \text{path}_j(q)$ the condition

$$\text{if} \quad n > 1 \quad \text{then} \quad |\gamma|_2 \geqq 1$$

holds, since from $|\gamma|_2 = 0$ it would follow that $\mathfrak{A}$ is a deleting DR transducer. Our first Lemma is essentially a consequence of this observation.

Lemma 1. Let $\mathfrak{A} = (F, A, G, P, a_0)$ be a UNDR transducer, moreover, $m \geqq 0$, $p \in T_{F,m}$, $q \in T_{G,m}$ and $a \in A$ be such that $(p, q) \in \tau'_{\mathfrak{A}(a)}$. Then

(a) for each $j \in [m]$ and $\alpha \in \text{path}_j(p)$ there exists a $\beta \in \text{path}_j(q)$ for which $|\alpha|_2 \leqq |\beta|_2$ and

(b) for each $j \in [m]$ and $\alpha \in \text{path}_j(q)$ there exists a $\beta \in \text{path}_j(p)$ with $|\beta|_2 \leqq |\alpha|_2$.

Proof. We prove only the part (a) of our statement since (b), as a converse of (a), can be shown in a similar way. We follow an induction on p.

If $p = x_i$ for some $i \in [m]$ then $q = x_i$ hence (a) trivially holds.

Now let $p = f(p_1, \ldots, p_n)$ for some $n \geqq 0$, $f \in F_n$ and $p_1, \ldots, p_n \in T_{F,m}$. By our supposition, there exists a rule $af(x_1, \ldots, x_n) \to \bar{q}(a_1 x_1, \ldots, a_n x_n)$ in P and there are $q_1, \ldots, q_n \in T_{G,m}$ for which $(p_i, q_i) \in \tau'_{\mathfrak{A}(a_i)}$ $(i \in [n])$ and $q = \bar{q}(q_1, \ldots, q_n)$. Since the case $n = 0$ is again trivial we may suppose that $n \geqq 1$. Then $\alpha = i\alpha'$ for some $i \in [n]$ and $\alpha' \in \text{path}_j(p_i)$.

Let γ be an arbitrary element of $\text{path}_i(\bar{q})$, which is not empty since $\mathfrak{A}$ is non-deleting, and let $\beta' \in \text{path}_j(q_i)$ be such that $|\alpha'|_2 \leqq |\beta'|_2$. Put $\beta = \gamma\beta'$. We mention that β' exists because of the induction hypothesis and, obviously, $\beta \in \text{path}_j(q)$. Now we distinguish the cases $n = 1$ and $n > 1$.

If $n = 1$ then

$$|\alpha|_2 = |i\alpha'|_2 = |\alpha'|_2 \leqq |\beta'|_2 \leqq |\gamma|_2 + |\beta'|_2 = |\gamma\beta'|_2 = |\beta|_2.$$

On the other hand, in the case $n > 1$, by our above note on $|\gamma|_2$ we have

$$|\alpha|_2 = |i\alpha'|_2 = 1 + |\alpha'|_2 \leqq |\gamma|_2 + |\beta'|_2 = |\gamma\beta'|_2 = |\beta|_2.$$

The proof is complete. □

Now we recall what we mean by the syntactic composition of two DR transducers. The exact definition can be found in [1].

Let $\mathfrak{A}_1 = (F^0, A_1, F^1, P_1, a_1)$ and $\mathfrak{A}_2 = (F^1, A_2, F^2, P_2, a_2)$ be two arbitrary DR transducers. Their syntactic composition is the DR transducer $\mathfrak{A}_1 \circ \mathfrak{A}_2 = (F^0, A_1 \times A_2, F^2, P, (a_1, a_2))$ where P is constructed in the following way. Whenever $bf(x_1, \ldots, x_n) \to \bar{q}(b_1 x_{i_1}, \ldots, b_m x_{i_m})$ is a rule in P_1 and it holds that $c\bar{q} \overset{*}{\underset{\mathfrak{A}_2}{\Rightarrow}} q(c_{1_1}x_1, \ldots, c_{1_{v_1}}x_1, \ldots, c_{m_1}x_m, \ldots, c_{m_{v_m}}x_m)$ we put the rule $(b, c)f(x_1, \ldots, x_n) \to q((b_1, c_{1_1})x_{i_1}, \ldots, (b_1, c_{1_{v_1}})x_{i_1}, \ldots, (b_m, c_{m_1})x_{i_m}, \ldots, (b_m, c_{m_{v_m}})x_{i_m})$ in P. It is well known that this construction yields the application of $\mathfrak{A}_2$ after $\mathfrak{A}_1$ in a "step by step" way. A very useful property of the syntactic composition is the following: if $\mathfrak{A}_2$ is an NDR transducer then $\tau_{\mathfrak{A}_1 \circ \mathfrak{A}_2} = \tau_{\mathfrak{A}_1} \circ \tau_{\mathfrak{A}_2}$ (for a proof, see Lemma 3 in [1]).

We shall need the generalisation of the syntactic composition and the above equality for any $m \geqq 2$. Therefore we make the following definition.

Definition 1. Let $m \geqq 2$ and let $\mathfrak{A}_i$ be a DR transducer for each $i \in [m]$. By the syntactic composition of $\mathfrak{A}_1, \ldots, \mathfrak{A}_m$ we mean the DR transducer defined above if $m=2$ and the DR transducer $(\mathfrak{A}_1 \circ \ldots \circ \mathfrak{A}_{m-1}) \circ \mathfrak{A}_m$ if $m > 2$.

Then, using Lemma 3 of [1] as a basis, the following statement can be verified by an induction on m: if $\mathfrak{A}_2, \ldots, \mathfrak{A}_m$ are NDR transducers then $\tau_{\mathfrak{A}_1 \circ \ldots \circ \mathfrak{A}_m} = \tau_{\mathfrak{A}_1} \circ \ldots \circ \tau_{\mathfrak{A}_m}$.

The next lemma says that this equality is also valid for extended tree transformations.

Lemma 2. Let $m \geqq 2$ and let $\mathfrak{A}_i = (F^{i-1}, A_i, F^i, P_i, a_i)$ be a DR transducer for each $i \in [m]$ such that $\mathfrak{A}_2, \ldots, \mathfrak{A}_m$ are NDR transducers. Then $\tau'_{\mathfrak{A}_1 \circ \ldots \circ \mathfrak{A}_m} = \tau'_{\mathfrak{A}_1} \circ \ldots \circ \tau'_{\mathfrak{A}_m}$.

Proof. Induction on m. For $m=2$ it is enough to show that for each $n \geqq 0$, $p \in T_{F^0,n}$, $q \in T_{F^2,n}$, $b_1 \in A_1$ and $b_2 \in A_2$ the following equivalence holds:

$$(p, q) \in \tau'_{\mathfrak{A}_1 \circ \mathfrak{A}_2((b_1, b_2))} \Leftrightarrow (\exists r \in T_{F^1,n})((p, r) \in \tau'_{\mathfrak{A}_1(b_1)} \text{ and } (r, q) \in \tau'_{\mathfrak{A}_2(b_2)}).$$

This can be verified by an induction on p. The detailed proof is omitted.

Finally, the induction step of m is shown by the following computation

$$\tau'_{\mathfrak{A}_1 \circ \ldots \circ \mathfrak{A}_m} = \tau'_{(\mathfrak{A}_1 \circ \ldots \circ \mathfrak{A}_{m-1}) \circ \mathfrak{A}_m} = \tau'_{\mathfrak{A}_1 \circ \ldots \circ \mathfrak{A}_{m-1}} \circ \tau'_{\mathfrak{A}_m} = \tau'_{\mathfrak{A}_1} \circ \ldots \circ \tau'_{\mathfrak{A}_m}. \quad \square$$

At this point we declare our main theorem.

Theorem 3. For any $m \geqq 2$ and $1 \leqq k < m$ $(\mathcal{LNDR} \circ \mathcal{NH})^k \subset (\mathcal{LNDR} \circ \mathcal{NH})^m$.

Proof. Because the complete proof is rather long we structured it in the following way. First we give a tree transformation τ_m which is in $(\mathcal{LNDR} \circ \mathcal{NH})^m$. Then we present Lemma 4 which concers any $\mathcal{NDR}$ transducer which induces τ_m. After this we suppose that $\tau_m \in (\mathcal{LNDR} \circ \mathcal{NH})^k$ for some $k \leqq m$ and, during a series of lemmas from 5 to 14, show, in Lemma 14, that $k < m$ is impossible.

Take an arbitrary integer $m \geqq 2$ and keep it fixed in the rest of the proof of this theorem. To define τ_m we introduce an LNDR transducer $\mathfrak{A}$ and an NH transducer $\mathfrak{B}$ as follows.

Let the LNDR transducer $\mathfrak{A} = (F, \{a, d\}, F', P, a)$ be determined by the following conditions:

(a) $F = F_0 \cup F_2 \cup F_3$, $F_0 = \{\#\}$, $F_2 = \{f_2\}$ and $F_3 = \{g_3\}$;

(b) $F'=F'_0\cup F'_2\cup F'_3$, $F'_0=\{\#\}$, $F'_2=\{f_2,f'_2\}$ and $F'_3=\{g_3\}$;

(c) P consists of the rules (i)—(vi) listed below

(i) $a\# \to \#$

(ii) $af_2(x_1, x_2)\to f'_2(dx_1, dx_2)$

(iii) $ag_3(x_1, x_2, x_3)\to g_3(dx_1, ax_2, dx_3)$

(iv) $d\# \to \#$

(v) $df_2(x_1, x_2)\to f_2(dx_1, dx_2)$

(vi) $dg_3(x_1, x_2, x_3)\to g_3(dx_1, dx_2, dx_3)$.

Moreover, introduce the NH transducer $\mathfrak{B}=(F', \{b\}, F, P', b)$ with P' containing the following rules:

(i) $b\# \to \#$

(ii) $bf'_2(x_1, x_2)\to g_3(bx_1, bx_1, bx_2)$

(iii) $bf_2(x_1, x_2)\to f_2(bx_1, bx_2)$

(iv) $bg_3(x_1, x_2, x_3)\to g_3(bx_1, bx_2, bx_3)$.

It is not difficult to see how $\mathfrak{A}$ and after that $\mathfrak{B}$ works on a tree $p\in T_F$. First $\mathfrak{A}$, with its state $a\in A$, searches for the first occurrence of f_2 on the path of p leading along the "middle branches" of a (possibly empty) sequence of g_3's and if it is found then rewrites it to f'_2 producing a tree $p'\in T_{F'}$. Any other symbol of p stays as it was. Then $\mathfrak{B}$ looks for this f'_2 in p' and duplicates the subtree on the first branch of f'_2 by substituting f'_2 by g_3. The other symbols of p' remain unchanged.

We put $\tau_m=(\tau_{\mathfrak{A}}\circ\tau_{\mathfrak{B}})^m$. Of course, $\tau_m\in(\mathcal{L}\mathcal{N}\mathcal{D}\mathcal{R}\circ\mathcal{N}\mathcal{H})^m$.

Now, for each $i\geqq 1$, we define a pair of trees $P_i, Q_i\in T_{F,i+1}$ recursively as follows:

(a) $P_1=f_2(x_2,x_1)$, $Q_1=g_3(x_2, x_2, x_1)$,

(b) $P_i=f_2(P_{i-1}(x_2, \ldots, x_{i+1}), x_1)$, $Q_i=g_3(P_{i-1}(x_2, \ldots, x_{i+1}), Q_{i-1}(x_2, \ldots, x_{i+1}), x_1)$ if $i>1$.

To make it clearer, P_m and Q_m are visualized in Fig. 1.

Let us introduce the notation $\tau'_m=(\tau'_{\mathfrak{A}}\circ\tau'_{\mathfrak{B}})^m$. By the definitions of $\mathfrak{A}$ and $\mathfrak{B}$, it can easily be verified that $(P_m, Q_m)\in\tau'_m$ moreover, that for each $t_1, \ldots, t_{m+1}\in T_F$ it holds

$$(P_m(t_1, \ldots, t_{m+1}), Q_m(t_1, \ldots, t_{m+1}))\in\tau_m. \tag{5}$$

Lemma 4. Let $\mathfrak{L}=(F, C, F, P'', c_0)$ be an NDR transducer with $\tau_{\mathfrak{L}}=\tau_m$. Then we have $(P_m, Q_m)\in\tau'_{\mathfrak{L}}$.

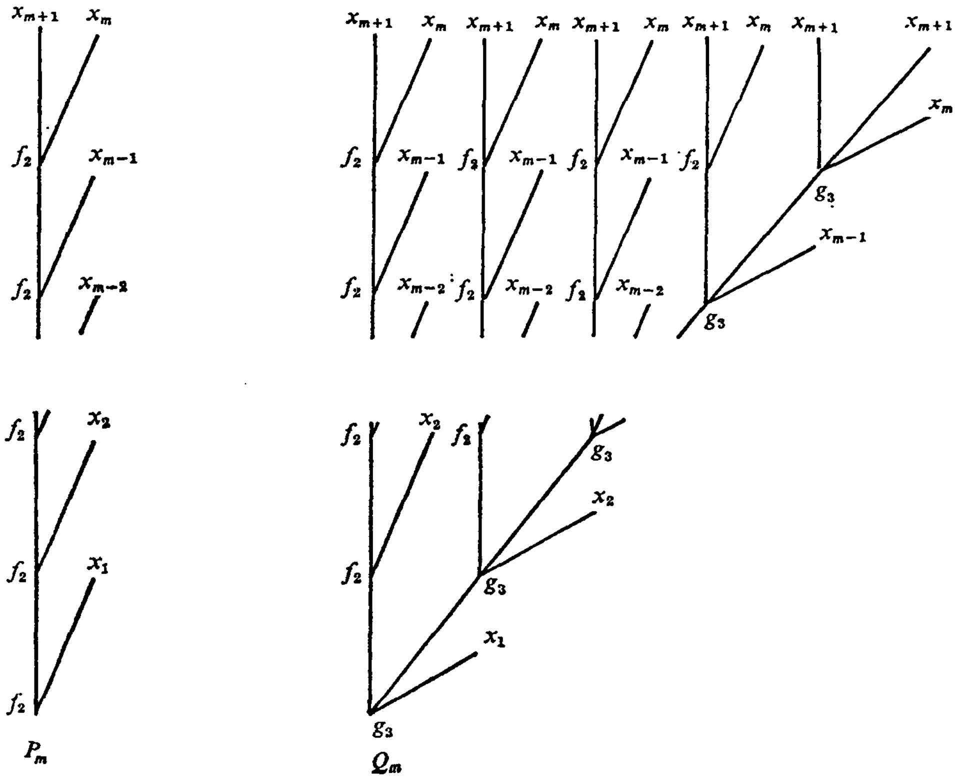

Figure 1.

Proof. First we note that $P_m \in \operatorname{dom} \tau'_{\mathfrak{L}}$ since, among others,

$$P_m(\#, ..., \#) \in \operatorname{dom} \tau_m = \operatorname{dom} \tau_{\mathfrak{L}}$$

thus (P_m, R_m) must be in $\tau'_{\mathfrak{L}}$ for some $R_m \in T_{F,m+1}$. It is obvious that R_m can be written in the form $\bar{R}_m(\overbrace{x_1, ..., x_1}^{n_1 \text{ times}}, ..., \overbrace{x_{m+1}, ..., x_{m+1}}^{n_{m+1} \text{ times}})$ for some $n_i \geqq 1$ $(i \in [m+1])$, $\bar{R}_m \in \hat{T}_{F,n}$ where $n = n_1 + ... + n_{m+1}$. Then it follows that for each $t_1, ..., t_{m+1} \in T_F$ the tree $\tau_{\mathfrak{L}}(P_m(t_1, ..., t_{m+1}))$ can be written in the form $\bar{R}_m(t_{1_1}, ..., t_{1_{n_1}}, ..., t_{m+1_1}, ..., t_{m+1_{n_{m+1}}})$ where for each $i \in [m+1]$ and $j \in [n_i]$ $(t_i\, t_{i_j}) \in \tau_{\mathfrak{L}(c_{i_j})}$ for some $c_{i_j} \in C$. Using these notations we have that for each $t_1, ..., t_{m+1} \in T_F$

$$Q_m(t_1, ..., t_{m+1}) = \bar{R}_m(t_{1_1}, ..., t_{1_{n_1}}, ..., t_{m+1_1}, ..., t_{m+1_{n_{m+1}}}).$$

We shall use this equality under different choices of $t_1, ..., t_{m+1}$ in the sequel of this proof.

Now suppose that $Q_m \neq R_m$. This means that for some $\alpha \in \operatorname{path}(Q_m) \cap \cap \operatorname{path}(R_m)$, $\operatorname{lab}(Q_m, \alpha) \neq \operatorname{lab}(R_m, \alpha)$. Then four different cases are possible, each of which yields a contradiction:

(a) $\mathrm{lab}(Q_m, \alpha)=f$, $\mathrm{lab}(R_m, \alpha)=g$ for some $f, g\in F$ such that $f\neq g$. But then

$$f=\mathrm{lab}(Q_m, \alpha)=\mathrm{lab}(Q_m(\#, \dots, \#), \alpha)=\mathrm{lab}(\bar{R}_m(\#_{1_1}, \dots, \#_{1_{n_1}}, \dots; \#_{m+1_1}, \dots$$
$$\dots, \#_{m+1_{n_{m+1}}}), \alpha) = \mathrm{lab}(R_m, \alpha)=g$$

which is impossible.

(b) $\mathrm{lab}(Q_m, \alpha)=x_i$, $\mathrm{lab}(R_m, \alpha)=g$ for some $i\in[m+1]$ and $g\in F$. Now, on the one hand $g=\#$, by $\#=\mathrm{lab}(Q_m(\#, \dots, \#), \alpha)=\mathrm{lab}(\bar{R}_m(\#_{1_1}, \dots, \#_{1_{n_1}}, \dots, \#_{m+1_1}, \dots, \#_{m+1_{n_{m+1}}}), \alpha)=\mathrm{lab}(R_m, \alpha)=g$. On the other hand, for any $t\in T_F$

$$t = \mathrm{lab}(Q_m(\#, \dots, \overset{i\text{-th}}{t}, \dots, \#), \alpha) =$$
$$= \mathrm{lab}(\bar{R}_m(\#_{1_1}, \dots, \#_{1_{n_1}}, \dots, t_{i_1}, \dots, t_{n_i}, \dots, \#_{m+1_1}, \dots, \#_{m+1_{n_{m+1}}}), \alpha) =$$
$$= \mathrm{lab}(R_m, \alpha) = g = \#,$$

a contradiction.

(c) $\mathrm{lab}(Q_m, \alpha)=f$, $\mathrm{lab}(R_m, \alpha)=x_i$ for some $f\in F$ and $i\in[m+1]$. Then it can be seen from the definition of Q_m (see Fig. 1) that in this case $\mathrm{str}(Q_m, \alpha)$ contains at least one x_j with $j\neq i$ whatever i be. But then for any $t\in T_F$ it holds that t is a subtree of $\mathrm{str}(Q_m(\#, \dots, \overset{j\text{-th}}{t}, \dots, \#), \alpha)$. It also holds that

$$\mathrm{str}(Q_m, (\#, \dots, \overset{j\text{-th}}{t}, \dots, \#), \alpha) =$$
$$= \mathrm{str}(\bar{R}_m(\#_{1_1}, \dots, \#_{1_{n_1}}, \dots, t_{j_1}, \dots, t_{j_{n_j}}, \dots, \#_{m+1_1}, \dots, \#_{m+1_{n_{m+1}}}), \alpha) = \#_{i_l}$$

for some $l\in[n_i]$. Contradiction since $\#_{i_l}$ does not depend on t chosen arbitrarily.

(d) $\mathrm{lab}(Q_m, \alpha)=x_i$, $\mathrm{lab}(R_m, \alpha)=x_j$ for some $i, j\in[m+1]$ such that $i\neq j$. Let t be an arbitrary element of T_F with $\mathrm{rn}(t)>0$. We have that

$$\# = \mathrm{lab}(Q_m(\#, \dots, \overset{j\text{-th}}{t}, \dots, \#), \alpha) =$$
$$= \mathrm{lab}(\bar{R}_m(\#_{1_1}, \dots, \#_{1_{n_1}}, \dots, t_{j_1}, \dots, t_{j_{n_j}}, \dots, \#_{m+1_1}, \dots, \#_{m+1_{n_{m+1}}}), \alpha) = t_{j_l}$$

for some $l\in[n_j]$, moreover $(t, t_{j_l})\in\tau_{\mathfrak{L}(c_{j_l})}$ for some $c_{j_l}\in C$. But $\mathfrak{L}$ is an NDR transducer hence $\mathrm{rn}(t_{j_l})>0$ which is again impossible. □

Let $k\leq m$ and assume that $\tau_m\in(\mathscr{LNDR}\circ\mathscr{NH})^k$. This means in a more detailed form that for each $i\in[2k]$ there exists a DR transducer $\mathfrak{A}_i=(F^{i-1}, A_i, F^i, P_i, a_i)$ such that the following conditions hold

(a) $F^0=F^{2k}=F$,

(b) if i is odd then $\mathfrak{A}_i$ is an LNDR transducer (6)

(c) if i is even then $\mathfrak{A}_i$ is an NH transducer

(d) $\tau_{\mathfrak{A}_1}\circ\dots\circ\tau_{\mathfrak{A}_{2k}}=\tau_m$.

Since each $\mathfrak{A}_i$ is an NDR transducer too, combining Lemmas 2 and 4 we have that

$$(P_m, Q_m) \in \tau'_{\mathfrak{A}_1} \circ \ldots \circ \tau'_{\mathfrak{A}_{2k}}, \tag{7}$$

or in other words, for each $i \in [2k]$ there exists a pair of trees $r_{i-1} \in T_{F^{i-1}, m+1}$ and $r_i \in T_{F^i, m+1}$ for which $r_0 = P_m$, $r_{2k} = Q_m$ and $(r_{i-1}, r_i) \in \tau'_{\mathfrak{A}_i}$. In fact, (7) is the relation which leads us to a contradiction in Lemma 14.

Lemma 5. For each $i \in [2k]$ it holds that $\mathrm{rn}_0(r_{i-1}) = 0$, that is, r_{i-1} does not contain any symbol of arity 0.

Proof. We observe that if $\mathrm{rn}_0(r_{i-1}) \neq 0$ then $\mathrm{rn}_0(r_i) \neq 0$ since $\mathfrak{A}_i$ is an NDR transducer. Now if for some $i \in [2k]$ $\mathrm{rn}_0(r_{i-1}) \neq 0$ then we obtain that $\mathrm{rn}_0(r_{2k}) \neq 0$ which, by the definition of r_{2k}, is a contradiction. □

Next we make a remark on the paths of r_0 and r_{2k} leading to x_j's $(j \in [m+1])$. Namely, we observe that whenever $j \in [m+1]$ and α is an element of either $\mathrm{path}_j(r_0)$ or $\mathrm{path}_j(r_{2k})$, by the definition of r_0 and r_{2k}, we have

$$|\alpha|_2 = \begin{cases} j & \text{if } j \in [m] \\ m & \text{if } j = m+1. \end{cases} \tag{8}$$

This can easily be read from Figure 1.

Lemma 6. For each $i \in [2k]$, $j \in [m+1]$ and $\alpha \in \mathrm{path}_j(r_{i-1})$, $|\alpha|_2$ is the same as in (8).

Proof. By part (a) of Lemma 1, for each $\alpha \in \mathrm{path}_j(r_{i-1})$ there exists a $\beta \in \mathrm{path}_j(r_i)$ with $|\alpha|_2 \leqq |\beta|_2$. Hence, if for some $i \in [2k]$ $j \in [m+1]$ and $\alpha \in \mathrm{path}_j(r_{i-1})$, $|\alpha|_2 > j$ when $j \in [m]$ and $|\alpha|_2 > m$ when $j = m+1$ holds then we obtain that for some $\beta \in \mathrm{path}_j(r_{2k})$, $|\beta|_2 > j$ if $j \in [m]$ and $|\beta|_2 > m$ if $j = m+1$. This, however, contradicts (8).

In a similar way, using part (b) of Lemma 1, the condition $|\alpha|_2 < j$ if $j \in [m]$ and $|\alpha|_2 < m$ if $j = m+1$ yields the existence of a $\beta \in \mathrm{path}_j(r_0)$ with the same property as α has, contradicting again (8). □

Lemma 7. Let $i \in [2k]$ and $r_i' \in T_{F^i}(A_i X_{m+1})$ be such that

$$a_i r_{i-1} \overset{*}{\underset{\mathfrak{A}_i}{\Rightarrow}} r_i'.$$

Suppose that the rule $cf(x_1, \ldots, x_n) \to q(c_1 x_1, \ldots, c_n x_n)$ was applied in the above derivation, where $f \in F_n^{i-1}$ for some $n \geqq 1$, $c, c_1, \ldots, c_n \in A_i$ and $q \in T_{F^i, n}$. Then for each $j \in [n]$ and $\gamma \in \mathrm{path}_j(q)$ it holds

$$|\gamma|_2 = \begin{cases} 0 & \text{if } n = 1, \\ 1 & \text{if } n > 1. \end{cases}$$

(We mention that r_i belongs to r_i'.)

Proof. By the conditions of our lemma, there exist the terms $s_{i-1} \in T_{F^{i-1}, m+2}$, $t_1, \ldots, t_n \in T_{F^{i-1}, m+1}$, $s_i \in T_{F^i, m+2}$ and $q_1, \ldots, q_n \in T_{F^i, m+1}$ such that each of the following conditions is satisfied.

(a) s_{i-1} contains exactly one occurrence of x_{m+2},

(b) $r_{i-1} = s_{i-1} \cdot_{m+2} f(t_1, ..., t_n)$,

(c) $r_i = s_i \cdot_{m+2} q(q_1, ..., q_n)$,

(d) $(s_{i-1}, s_i) \in \tau'_{\mathfrak{A}_i}$, $(t_j, q_j) \in \tau'_{\mathfrak{A}_i(c_j)}$ $(j \in [n])$.

Let us suppose that for some $j \in [n]$ and $\gamma \in \text{path}_j(q)$ $|\gamma|_2$ violates the condition stated by our lemma. By Lemma 5 and (a) we can choose an $l \in [m+1]$ such that for some $\alpha \in \text{path}_l(r_{i-1})$ α can be written in the form $\alpha = \alpha_1 j \alpha_2$ where $\alpha_1 \in \text{path}_{m+2}(s_{i-1})$ and $\alpha_2 \in \text{path}_l(t_j)$. Moreover, by Lemma 1, there exist $\beta_1 \in \text{path}_{m+2}(s_i)$ and $\beta_2 \in \text{path}_l(q_j)$ with $|\alpha_1|_2 \leqq |\beta_1|_2$ and $|\alpha_2|_2 \leqq |\beta_2|_2$. Letting $\beta = \beta_1 \gamma \beta_2$ we obviously have that $\beta \in \text{path}_l(r_i)$.

First consider the case $n=1$. By our indirect assumption, $|\gamma|_2 > 0$, from which we have

$$|\alpha|_2 = |\alpha_1 j \alpha_2|_2 = |\alpha_1|_2 + |\alpha_2|_2 < |\beta_1|_2 + |\gamma|_2 + |\beta_2|_2 = |\beta_1 \gamma \beta_2|_2 = |\beta|_2$$

contradicting Lemma 6.

Now assume that $n > 1$. In this case $|\gamma|_2 = 0$ is impossible by our observation made at the beginning of this section hence the indirect assumption is $|\gamma|_2 > 1$. But then

$$|\alpha|_2 = |\alpha_1 j \alpha_2|_2 = |\alpha_1|_2 + 1 + |\alpha_2|_2 < |\beta_1|_2 + |\gamma|_2 + |\beta_2|_2 = |\beta_1 \gamma \beta_2|_2 = |\beta|_2,$$

a contradiction. □

Lemma 8. For each $i \in [2k]$ and $n \geqq 4$, $\text{m}_n(r_{i-1}) = 0$.

Proof. Suppose it does not hold. Let $i \in [2k]$ be the greatest integer for which $\text{m}_n(r_{i-1}) > 0$ for some $n \geqq 4$. Then in the derivation $a_i r_{i-1} \overset{*}{\underset{\mathfrak{A}_i}{\Rightarrow}} r'_i \left(r'_i \in T_{F^i}(A_i X_{m+1})\right)$ it has to be applied at least one rule $cf(x_1, ..., x_n) \to q(cx_1, ..., cx_n)$ for which $n \geqq 4$ and $\text{m}_l(q) = 0$ for each $l \geqq 4$. Since $\mathfrak{A}_i$ is an NDR transducer it can be possible only if $|\gamma|_2 > 1$ for some $j \in [n]$ and $\gamma \in \text{path}_j(q)$. This is a contradiction, by Lemma 7. □

At this point of the proof we can declare that for each $i \in [2k]$, every function symbol of r_{i-1} is in $F_1^{i-1} \cup F_2^{i-1} \cup F_3^{i-1}$.

Lemma 9. Let $i \in [2k]$ and $r'_i \in T_{F^i}(A_i X_{m+1})$ be such that

$$a_i r_{i-1} \overset{*}{\underset{\mathfrak{A}_i}{\Rightarrow}} r'_i.$$

Suppose that in the above derivation it was applied a rule $cf(x_1, ..., x_n) \to q$ where $f \in F_n^{i-1}$, $n \geqq 1$, $c \in A_i$ and $q \in T_{F^i}(A_i X_n)$. Then $n \in [3]$ and q can be written in one of the following three forms, for some suitable $u_0, u_1, u_2, u_3 \in T_{F_1^i}$, $c_1, c_2, c_3 \in A_i$, $g \in F_2^i$ and $h \in F_3^i$:

(a) if $n = 1$ then $q = u_0(c_1 x_1)$,

(b) if $n = 2$ then either

$$q = u_0(g(u_1(c_1x_{i_1}), u_2(c_2x_{i_2}))) \text{ where } \{i_1, i_2\} = [2] \text{ or} \quad (9)$$

$$q = u_0(h(u_1(c_1x_{i_1}), u_2(c_2x_{i_2}), u_3(c_3x_{i_3}))) \text{ where}$$

$$\{i_1, i_2, i_3\} = [2],$$

(c) if $n = 3$ then

$$q = u_0(h(u_1(c_1x_{i_1}), u_2(c_2x_{i_2}), u_3(c_3x_{i_3}))) \text{ with}$$

$$\{i_1, i_2, i_3\} = [3].$$

(We note that in the notation $T_{F_1^i,1}$, F_1^i is considered a ranked alphabet. Thus the condition $u_j \in T_{F_1^i,1}$ means that every function symbol of u_j is in F_1^i ($j=0,1,2,3$).)

Proof. Immediate from Lemmas 5, 7, 8 and from the fact that $\mathfrak{A}_i$ is an NDR transducer. □

Definition. We say that for some $i\in[2k]$ r_{i-1} has property (10) if

(a) for some $f\in F_3^{i-1}$ and $p_1, p_2, p_3 \in T_{F^{i-1},m+1}$, $f(p_1, p_2, p_3)\in \text{sub}(r_{i-1})$

and (10)

(b) for each $j\in[3]$, $n_j > 0$ where $n_j = \max\{|\alpha|_2 \mid \alpha\in \text{path}_l(p_j),\ l\in[m+1]\}$.

Lemma 10. There exists no $i\in[2k]$ for which r_{i-1} has property (10).

Proof. It is enough to show that whenever r_{i-1} has property (10) then so does r_i. This proves our lemma since r_{2k}, by its definition, does not have property (10).

To this end, let us suppose that r_{i-1} has property (10) ($i\in[2k]$). Then, from Lemma 9, it follows that for some suitable $s_{i-1}\in T_{F^{i-1},m+2}$, $s_i\in T_{F^i,m+2}$, $u_0, u_1, u_2, u_3 \in T_{F_1^i,1}$, $c, c_1, c_2, c_3 \in A_i$ and $q_1, q_2, q_3 \in T_{F^i,m+1}$ the following relations hold:

(a) $r_{i-1} = s_{i-1} \cdot_{m+2} f(p_1, p_2, p_3)$,

(b) $r_i = s_i \cdot_{m+2} u_0(h(u_1(q_1), u_2(q_2), u_3(q_3)))$,

(c) $cf(x_1, x_2, x_3) \to u_0(h(u_1(c_1x_{i_1}), u_2(c_2x_{i_2}), u_3(c_3x_{i_3})))\in P_i$, $\{i_1, i_2, i_3\}=[3]$,

(d) $(s_{i-1}, s_i)\in\tau'_{\mathfrak{A}_i}$, $(p_{i_j}, q_j)\in\tau'_{\mathfrak{A}_i(c_j)}$ for each $j\in[3]$.

Moreover, for each $j\in[3]$, there exist $l_j\in[m+1]$ and $\alpha_j\in\text{path}_{l_j}(p_{i_j})$ with $|\alpha_j|_2=n_{i_j}$. By Lemma 1, there exist $\beta_j\in\text{path}_{l_j}(q_j)$ such that $|\alpha_j|_2\leq|\beta_j|_2$. This shows that r_i has property (10) with $h(u_1(q_1), u_2(q_2), u_3(q_3))$. □

Definition. Let $i\in[2k]$. We say that r_{i-1} has property (11) if

(a) for some $f\in F_3^{i-1}$ and $p_1, p_2, p_3\in T_{F^{i-1},m+1}$, $f(p_1, p_2, p_3)\in\text{sub}(r_{i-1})$

and (11)

(b) there exists exactly one $j\in[3]$ with $n_j>0$ where n_j is the same as in the definition of property (10).

Lemma 11. There exist no $i\in[2k]$ such that r_{i-1} has property (11).

Proof. Since r_{2k} does not have property (11), we can use the same technique as in the proof of Lemma 10. Assume that r_{i-1} has property (11). Then, using the notations of Lemma 10, we again have (a)—(d) as in Lemma 10 and, without loss of generality, may suppose that $|\alpha_1|_2>0$, $|\alpha_2|_2=|\alpha_3|_2=0$ or, in other words, $p_{i_2}, p_{i_3}\in T_{F_1^{i-1},m+1}$. Hence, from Lemmas 1 and 9 it follows that $|\beta_1|_2>0$ and $q_2, q_3\in T_{F_1^i,m+2}$ meaning that r_i has property (11). □

We shall need one further property.

Definition. We say that for some $i\in[2k]$, r_{i-1} has property (12) if there exist $\alpha, \beta\in\text{path}(r_{i-1})$ satisfying the following conditions:

(a) $\alpha \not\leqq \beta$ and $\beta \not\leqq \alpha$,

(b) $\text{str}(r_{i-1}, \alpha) = f(p_1, p_2, p_3)$ for some $f\in F_3^{i-1}$,

$p_1, p_2, p_3\in T_{F^{i-1},m+1}$, (12)

(c) $\text{str}(r_{i-1}, \beta) = f'(p_1', p_2', p_3')$ for some $f'\in F_3^{i-1}$,

$p_1', p_2', p_3'\in T_{F^{i-1},m+1}$.

Lemma 12. There exist no $i\in[2k]$ for which r_{i-1} has property (12).

Proof. If r_{i-1} has property (12) then, by Lemma 9, so does r_i. This proves our lemma since r_{2k} does not have property (12). □

Lemma 13. Let $i\in[2k]$ be an odd integer. Then $\text{rn}_3(r_{i-1})=\text{rn}_3(r_i)$.

Proof. It obviously follows from Lemma 9 that $\text{rn}_3(r_{i-1})\leqq\text{rn}_3(r_i)$. Let us assume that $\text{rn}_3(r_{i-1})<\text{rn}_3(r_i)$. Then in the derivation $a_i r_{i-1} \overset{*}{\underset{\mathfrak{A}_i}{\Rightarrow}} r_i'$ $(r_i'\in T_{F^i}(A_i X_{m+1}))$ it has to be applied at least one rule of the form (9). However, this is impossible since, for odd i, $\mathfrak{A}_i$ is an LNDR transducer. □

Lemma 14. $k=m$.

Proof. On the contrary, assume that $k<m$. Then, since $\text{rn}_3(r_0)=0$ and $\text{rn}_3(r_{2k})=m$, it follows from Lemma 13 that for some even integer $i\in[2k]$, $\text{rn}_3(r_{i-1})\leqq\leqq\text{rn}_3(r_i)-2$. It means that there exist $\alpha, \beta\in\text{path}(r_{i-1})$ such that $\alpha\neq\beta$, $\text{str}(r_{i-1}, \alpha)= =f(p_1, p_2)$, $\text{str}(r_{i-1}, \beta)=f'(p_1', p_2')$ for some $f, f'\in F_2^{i-1}$, $p_j, p_j'\in T_{F^{i-1},m+1}$ $(j\in[2])$ moreover, in the derivation $a_i r_{i-1} \overset{*}{\underset{\mathfrak{A}_i}{\Rightarrow}} r_i'$ $(r_i'\in T_{F^i}(A_i X_{m+1}))$ both f and f' were rewritten by applying a rule of the form (9).

First we claim that either $\alpha<\beta$ or $\beta<\alpha$. Really, from $\alpha\nless\beta$, $\beta\nless\alpha$ and $\alpha\neq\beta$ it would follow that r_i has property (12) contradicting Lemma 12.

Suppose that $\alpha<\beta$ and that the rule (9) was applied to rewrite f in the derivation $a_i r_{i-1} \overset{*}{\underset{\mathfrak{A}_i}{\Rightarrow}} r_i'$. Then, without loss of generality, we may assume that the following relations hold for some suitable $s_{i-1}, t_{i-1}\in T_{F^{i-1},m+2}$, $s_i\in T_{F^i,m+2}$ and $q_1, q_2, q_3\in T_{F^i,m+1}$:

(a) $r_{i-1} = s_{i-1} \cdot_{m+2} f(p_1, p_2)$,

(b) $p_1 = t_{i-1} \cdot_{m+2} f'(p_1', p_2')$,

(c) $r_i = s_i \cdot_{m+2} u_0(h(u_1(q_1), u_2(q_2), u_3(q_3)))$,

(d) $(s_{i-1}, s_i) \in \tau'_{\mathfrak{A}_i}$, $(p_{i_j}, q_j) \in \tau'_{\mathfrak{A}_i(c_j)}$ for each $j \in [3]$.

Let us introduce the following notations:

$$m_j = \max \{|\alpha|_2 | \alpha \in \text{path}_l(p_j) \text{ for some } l \in [m+1]\} \quad (j \in [2]),$$

$$n_j = \max \{|\alpha|_2 | \alpha \in \text{path}_l(q_j) \text{ for some } l \in [m+1]\} \quad (j \in [3]).$$

We know, by (b), that $m_1 > 0$. Moreover $m_2 = 0$ since from $m_2 > 0$ it would follow, by (d) and Lemma 1, that $n_1 > 0$, $n_2 > 0$ and $n_3 > 0$. This, however would mean that r_i has property (10) which is impossible by Lemma 10. Hence we have $p_2 \in T_{F_1^{i-1}, m+1}$.

We also know, by (9), that $\{i_1, i_2, i_3\} = [2]$ which means that $\mathfrak{A}_i$ duplicates either p_1 or p_2. We show that both cases are impossible.

First let us suppose that 1 appears once and 2 appears twice in the sequence i_1, i_2, i_3. Then we obtain, by Lemmas 1 and 9, that for exactly one $j \in [3]$, $n_j > 0$, contradicting Lemma 11.

Next assume that 1 appears twice and 2 appears once in the sequence i_1, i_2, i_3. But then, since f' was also rewritten by a rule of type (9) we have that r_i has property (12) yielding again a contradiction, by Lemma 12.

Hence we have $k = m$. □

With this we also completed the proof of Theorem 3. □

Now we present Theorem 3 in an alternative form. It is not difficult to see that $\mathcal{LNDR} \circ \mathcal{NH} = \mathcal{UNDR}$. Really, for any LNDR transducer $\mathfrak{A}$ and NH transducer $\mathfrak{B}$, by Lemma 3 of [1], $\tau_{\mathfrak{A} \circ \mathfrak{B}} = \tau_{\mathfrak{A}} \circ \tau_{\mathfrak{B}}$ and it can easily be verified that in this case $\mathfrak{A} \circ \mathfrak{B}$ is a UNDR transducer. Conversely, given a UNDR transducer $\mathfrak{L}$, with the help of the usual relabeling technique (see, for example, Lemma 3.1 in [2], pp. 155) we can construct an LNDR transducer $\mathfrak{A}$ and an NH transducer $\mathfrak{B}$ with $\tau_{\mathfrak{L}} = \tau_{\mathfrak{A}} \circ \tau_{\mathfrak{B}}$. Thus Theorem 3 can be given in the following form as well.

Theorem 15. For any $m \geqq 1$, $\mathcal{UNDR}^m \subset \mathcal{UNDR}^{m+1}$.

The first problem, presented at the beginning of this section is answered by

Theorem 16. $[S]$ is an infinite set.

Proof. Immediately follows from Theorem 3. □

Now we deal with the second problem. The following lemma can be proved.

Lemma 17. There exists an NDR transducer $\mathfrak{A} = (F, \{a, b\}, F', P, a)$ such that $\tau_{\mathfrak{A}} \notin (\mathcal{LNDR} \circ \mathcal{NH})^k$ for any $k \geqq 1$.

Proof. Let $\mathfrak{A}$ be determined by the following conditions:

(a) $F=F_0\cup F_2,\ F_0=\{\#\},\ F_2=\{f_2\}$,

(b) $F'=F'_0\cup F'_2\cup F'_3,\ F'_0=\{\#\},\ F'_2=\{f_2\},\ F'_3=\{g_3\}$,

(c) P is the set of the rules:

(i) $a\#\to\#,\ b\#\to\#,$

(ii) $af_2(x_1,x_2)\to g_3(bx_1,ax_1,bx_2),\ bf_2(x_1,x_2)\to f_2(bx_1,bx_2)$.

Let P_m and Q_m be defined in the same way as in the proof of Theorem 3. It is not difficult to verify that for each $m\geqq 1$, $(P_m, Q_m)\in\tau'_{\mathfrak{A}}$.

Moreover, let Q_m be written in the form

$$\bar{Q}_m(x_1,x_2,x_2,\dots,\overbrace{x_{m+1},\dots,x_{m+1}}^{m+1\text{ times}})\quad\text{where}\quad \bar{Q}_m\in\hat{T}_{F,n}\quad (n=1+2+\dots+m+1).$$

Then we can say that for any $t_1,\dots,t_{m+1}\in T_F$

$$(P_m(t_1,\dots,t_{m+1}),\ \bar{Q}_m(t_1,t_2,t_2,\dots,\overbrace{t_{m+1},\dots,t_{m+1}}^{m\text{ times}},t'_{m+1}))\in\tau_{\mathfrak{A}}$$

holds where $t'_{m+1}=\tau_{\mathfrak{A}}(t_{m+1})$. Using this notation, the following lemma can be proved in a similar way as Lemma 4. Therefore we omit the proof.

Lemma 18. If $\mathfrak{L}$ is an NDR transducer with $\tau_{\mathfrak{A}}=\tau_{\mathfrak{L}}$ then for each $m\geqq 1$ $(P_m, Q_m)\in\tau'_{\mathfrak{L}}$ holds. □

Now we can complete the proof of Lemma 17. Suppose that

$$\tau_{\mathfrak{A}}\in(\mathscr{L}\mathscr{N}\mathscr{D}\mathscr{R}\circ\mathscr{N}\mathscr{H})^k$$

for some $k\geqq 1$. Then for each $i\in[2k]$ there exists a DR transducer $\mathfrak{A}_i=(F^{i-1}, A_i, F^i, P_i, a_i)$ with properties (a)—(c) of (6) and $\tau_{\mathfrak{A}}=\tau_{\mathfrak{A}_1}\circ\dots\circ\tau_{\mathfrak{A}_{2k}}$. Let m be chosen such that $k<m$. It follows from Lemmas 2 and 18 that $(P_m, Q_m)\in\tau'_{\mathfrak{A}_1}\circ\dots\circ\tau'_{\mathfrak{A}_{2k}}$. However, if we follow the proof of Theorem 3 from (7) then we see in Lemma 14 that this is a contradiction. This ends the proof of Lemma 17. □

The last theorem is an immediate consequence of Lemma 17.

Theorem 19. $\bigcup_{k=1}^{\infty}(\mathscr{L}\mathscr{N}\mathscr{D}\mathscr{R}\circ\mathscr{N}\mathscr{H})^k\subset\mathscr{N}\mathscr{D}\mathscr{R}$.

RESEARCH GROUP ON THEORY OF AUTOMATA
HUNGARIAN ACADEMY OF SCIENCES
SOMOGYI B. U. 7.
SZEGED, HUNGARY
H—6720

References

[1] FÜLÖP, Z. and S. VÁGVÖLGYI, Results on compositions of deterministic root-to-frontier tree transformations, Acta Cybernetica, v. 8, f. 1, 1987, 49—61.

[2] GÉCSEG, F. and M. STEINBY, Tree Automata, Akadémiai Kiadó, Budapest, 1984.

(Received Sept. 28, 1986)