Discussiones Mathematicae
General Algebra and Applications 21 (2001) 219–227

TREE TRANSFORMATIONS DEFINED BY HYPERSUBSTITUTIONS*

Sr. Arworn
Chiang Mai University, Department of Mathematics
50200 Chiang Mai, Thailand
e-mail: scmti002@cmu.chiangmai.ac.th

and

Klaus Denecke
University of Potsdam, Institute of Mathematics
14415 Potsdam, Germany
e-mail: kdenecke@rz.uni-potsdam.de

Abstract

Tree transducers are systems which transform trees into trees just as automata transform strings into strings. They produce transformations, i.e. sets consisting of pairs of trees where the first components are trees belonging to a first language and the second components belong to a second language. In this paper we consider hypersubstitutions, i.e. mappings which map operation symbols of the first language into terms of the second one and tree transformations defined by such hypersubstitutions. We prove that the set of all tree transformations which are defined by hypersubstitutions of a given type forms a monoid with respect to the composition of binary relations which is isomorphic to the monoid of all hypersubstitutions of this type. We characterize transitivity, reflexivity and symmetry of tree transformations by properties of the corresponding hypersubstitutions. The results will be applied to languages built up by individual variables and one operation symbol of arity $n \geq 2$.

Keywords: hypersubstitution, tree transformation, tree transducer.

2000 Mathematics Subject Classification: 08B15, 08B25.

*The research was supported by grants of the NRCT (Thailand) and the DFG (Germany).

1. Introduction

Let $\Sigma := \{f_i \mid i \in I\}$ be a set of operation symbols of type $\tau_1 = (n_i)_{i\in I}$, where f_i is n_i-ary, $n_i \in I\!N$ and let $\Omega = \{g_j \mid j \in J\}$ be a set of operation symbols of type $\tau_2 = (n_j)_{j\in J}$ where g_j is n_j-ary. We denote by $W_{\tau_1}(X)$ and by $W_{\tau_2}(X)$ the sets of all terms of type τ_1 and of type τ_2, respectively. Then we define

Definition 1.1. A $(\tau_1 - \tau_2)$-*hypersubstitution* is a mapping

$$\sigma : \{f_i \mid i \in I\} \to W_{\tau_2}(X),$$

which maps operation symbols of type τ_1 to terms of type τ_2 and preserves the arities.

Clearly, every $(\tau_1 - \tau_2)$-hypersubstitution σ can be extended to a mapping

$$\hat{\sigma} : W_{\tau_1}(X) \to W_{\tau_2}(X)$$

in the following inductive way:

(i) $\hat{\sigma}[x] := x$,

(ii) $\hat{\sigma}[f_i(t_i, \ldots, t_{n_i})] := \sigma(f_i)(\hat{\sigma}[t_i], \ldots, \hat{\sigma}[t_{n_i}])$.

Remark that the right hand side of (ii) means that one has to substitute the term $\hat{\sigma}[t_i]$ for the variables x_i, for $1 \le i \le n_i$, in the term $\sigma(f_i)$.

Definition 1.2. Let σ be a $(\tau_1 - \tau_2)$-hypersubstitution. Then

$$T_\sigma := \{(t, \hat{\sigma}[t]) \mid t \in W_{\tau_1}(X)\}$$

is called *tree transformation defined by* σ.

We remark that terms also are called *trees*.

Instead of the two different types τ_1 and τ_2 one can consider the join $\Sigma \cup \Omega$ of the operation symbols of type τ_1 and of type τ_2 to obtain a type τ. Then a *hypersubstitution of type* τ is a mapping

$$\sigma : \{h_l \mid l \in L\} \to W_\tau(X),$$

where $L = I \cup J$, which preserves the arities. Let $Hyp(\tau)$ be the set of all such hypersubstitutions. On $Hyp(\tau)$ a binary operation $\circ_h$ can be defined by

$$\sigma_1 \circ_h \sigma_2 := \hat{\sigma}_1 \circ \sigma_2,$$

where $\circ$ is the usual composition of functions.

Clearly, the hypersubstitution σ_{id} which maps each operation symbol h_l to the so-called fundamental term $h_l(x_1, \ldots, x_{n_l})$ is an identity element with respect to $\circ_h$ and $(Hyp(\tau); \circ_h, \sigma_{id})$ is a monoid. Regarding the type τ as the union of the types τ_1 and τ_2, every $(\tau_1 - \tau_2)$-hypersubstitution can be considered as a hypersubstitution of type τ, where the operation symbols g_j of the type τ_2 are mapped to the fundamental terms $g_j(x_1, \ldots, x_{n_j})$. Since the composition of two hypersubstitutions, which fix the operation symbols from Ω, is again a hypersubstitution which fixes the operation symbols from Ω, the set of all $(\tau_1 - \tau_2)$-hypersubstitutions regarded as hypersubstitutions of the type τ forms a submonoid of the monoid $Hyp(\tau)$. This allows us to consider tree transformations T_σ where $\sigma \in Hyp(\tau)$, $t \in W_\tau(X)$ and $\hat{\sigma}[t] \in W_\tau(X)$.

Hypersubstitutions were introduced in [2] to make the concept of a hyperidentity more precise. Let V be a variety of type τ, i.e. an equationally defined class of algebras of the same type τ. Then an equation $s \approx t$ of terms of type τ is said to be satisfied as a *hyperidentity* in V if for all $\sigma \in Hyp(\tau)$ the equations $\hat{\sigma}[s] \approx \hat{\sigma}[t]$ are identities in the variety V. A variety V is called *solid* if each of its identities is satisfied as a hyperidentity in V. All solid varieties of a given type τ form a complete sublattice of the lattice of all varieties of type τ. For more background see [4].

2. TREE TRANSFORMATIONS AND TREE TRANSDUCERS

The most important tree transformations are those which can be given in an effective way. Next we define a general system, called a *tree transducer*, which induces tree transformations.

The concept of a tree transducer which will be used in this paper is due to Thatcher ([6]).

Definition 2.1. A $(\tau_1 - \tau_2)$-*tree transducer* is a sequence $\mathcal{A} = (\Sigma, X, A, \Omega, P, A')$, where $\Sigma = \{f_i \mid i \in I\}$ is a finite set of operation symbols of type τ_1, X is a finite set of variables, $\Omega = \{g_j \mid j \in J\}$ is a finite set of operation symbols of type τ_2, $A = \{a_1, \ldots, a_m\}$ is a finite set of unary operation symbols, $A' \subseteq A$, and P is a finite set of *productions (rules of derivation)* of the forms

(i) $x \mapsto at$ for $x \in X,\ a \in A,\ t \in W_{\tau_2}(X)$,

(ii) $f_i(a_1\xi_1, \ldots, a_{n_i}\xi_{n_i}) \ \mapsto \ at(\xi_1, \ldots, \xi_{n_i})$, for $f_i \in \Sigma, a_1, \ldots, a_{n_i} \in A$, $a \in A,\ \xi_1, \ldots, \xi_{n_i} \in \chi_{n_i}$, where $\chi_{n_i} = \{\xi_1, \ldots, \xi_{n_i}\}$ is an auxiliary alphabet, and $t(\xi_1, \ldots, \xi_{n_i}) \in W_\Omega(X \cup \chi_{n_i})$.

Further we define for trees s, t: s *directly derives* t *in* $\mathcal{A}$, if t can be obtained from s by one of the following steps:

(1) replacing an occurrence of a variable $x \in X$ in s by the right hand side of a production of the form (i)

or

(2) replacing an occurrence of a subtree $f_i(a_1q_1, \ldots, a_{n_i}q_{n_i})$, where $a_1, \ldots, a_{n_i} \in A,\ q_1, \ldots, q_{n_i} \in W_{\tau_2}(X \cup \chi_{n_i})$, in s by $at(q_1, \ldots, q_{n_i})$ if $f_i(a_i\xi_1, \ldots, a_{n_i}\xi_{n_i}) \mapsto at(\xi_1, \ldots, \xi_{n_i})$ is a production.

If s directly derives t in $\mathcal{A}$, we write $s \rightarrow_{\mathcal{A}} t$.

s *derives* t *in* $\mathcal{A}$, if there is a sequence $s \rightarrow_{\mathcal{A}} s_1 \rightarrow_{\mathcal{A}} s_2 \rightarrow_{\mathcal{A}} \ldots \rightarrow_{\mathcal{A}} s_n = t$ of direct derivations of t from s or if $s = t$. In this case we write $s \Rightarrow^*_{\mathcal{A}} t$. If we regard $\rightarrow_{\mathcal{A}}$ and $\Rightarrow^*_{\mathcal{A}}$ as binary relations, then $\Rightarrow^*_{\mathcal{A}}$ is the reflexive and transitive closure of $\rightarrow_{\mathcal{A}}$.

Tree transformations induced by tree transducers can be defined in the following way:

Definition 2.2. If $\mathcal{A}$ is a $(\tau_1 - \tau_2)$-tree transducer, then

$$T_{\mathcal{A}} := \{(s,t) \mid s \in W_{\tau_1}(X), t \in W_{\tau_2}(X) \ \text{ and } \ (\exists a_0 \in A')[s \Rightarrow^*_{\mathcal{A}} a_0t]\}$$

is called *tree transformation induced by* $\mathcal{A}$.

(That means, tree transformations of the form $T_{\mathcal{A}}$ can be described in an effective (algorithmic) way).

It turns out that tree transformations T_σ for a hypersubstitution σ have this property, i.e. we have

Proposition 2.3. *If σ is a $(\tau_1 - \tau_2)$-hypersubstitution then one can define a tree transducer $\mathcal{A} = (\Sigma, X, A, \Omega, P, A')$ with $A = A' = \{a\}$ and $P = \{x \rightarrow_{\mathcal{A}} ax \mid x \in X\} \cup \{f_i(a\xi_1, \ldots, a\xi_{n_i}) \rightarrow_{\mathcal{A}} a\sigma(f_i)(\xi_1, \ldots, \xi_{n_i}) \mid i \in I\}$, such that $T_{\mathcal{A}} = T_\sigma$.*

Proof. Because of the definition of $T_{\mathcal{A}}$ and T_σ for abitrary terms $t \in W_{\tau_1}(X), t' \in W_{\tau_2}(X)$, we have to prove

$$\hat{\sigma}[t] = t' \Leftrightarrow t \Rightarrow^*_{\mathcal{A}} a_0 t'.$$

We proceed by induction on the complexity of the term t. For the base case, let $t = x$ be an element of X. Then $\hat{\sigma}[x] = x$, i.e. $t' = x$ and $x \Rightarrow^*_{\mathcal{A}} ax$, since $x \rightarrow_{\mathcal{A}} ax$ is a rule and conversely.

Assume now that $t = f_i(t_1, \ldots, t_{n_i})$ and assume that

$$\hat{\sigma}[t_j] = t'_j \Leftrightarrow t_j \Rightarrow^*_{\mathcal{A}} at'_j \text{ for all } j = 1, \ldots, n_i.$$

Then $\hat{\sigma}[t] = \sigma(f_i)(\hat{\sigma}[t_1], \ldots, \hat{\sigma}[t_{n_i}]) = t'$ and we have to find a derivation of t' from t. In fact, we have

$$\begin{aligned} t = f_i(t_1, \ldots, t_{n_i}) \quad & \Rightarrow^*_{\mathcal{A}} \quad f_i(a\hat{\sigma}[t_1], \ldots, a\hat{\sigma}[t_{n_i}]) \text{ by hypothesis of the induction,} \\ & \rightarrow_{\mathcal{A}} \quad a\sigma(f_i)(\hat{\sigma}[t_1], \ldots, \hat{\sigma}[t_{n_i}]) = a\hat{\sigma}[t] \text{ by the second rule.} \end{aligned}$$

Conversely, assume that there is a derivation of at' from t, i.e. $t \Rightarrow^*_{\mathcal{A}} at'$. Since $t = f_i(t_1, \ldots, t_{n_i})$ in the first step of the derivation only a rule of the second kind is applicable and therefore $t_j = at'_j$ for $j = 1, \ldots, n_i$ and

$$t \Rightarrow^*_{\mathcal{A}} f_i(at'_1, \ldots, at'_{n_i}) \Rightarrow^*_{\mathcal{A}} a\sigma(f_i)(t'_1, \ldots, t'_{n_i}).$$

But $t_j = at'_j$ means $t_j \Rightarrow^*_{\mathcal{A}} at'_j$, for $j = 1, \ldots, n_i$. By hypothesis of the induction we have $\hat{\sigma}[t_j] = t'_j$ and then

$$t \Rightarrow^*_{\mathcal{A}} a\sigma(f_i)(\hat{\sigma}[t_1], \ldots, \hat{\sigma}[t_{n_i}]) = a\hat{\sigma}[t] = at'.$$

Since this derivation is unique, we get $t' = \hat{\sigma}[t]$. ∎

Note that tree transducers considered in Proposition 2.3 are particular cases of so-called *HF-transducers* introduced e.g. in [5] and that hypersubstitutions are special cases of *tree homomorphisms* ([5]).

3. OPERATIONS DEFINED ON TREE TRANSFORMATIONS

In the sequel we assume that we have only one type τ. By $T_{\sigma_1} \circ T_{\sigma_2}$ we denote the composition of the tree transformations T_{σ_1} and T_{σ_2}. Since a tree transformation T_σ is a relation, we can consider inverses, domains and ranges of such transformations under the relational composition $\circ$. We define $T_{Hyp(\tau)} := \{T_\sigma \mid \sigma \in Hyp(\tau)\}$ and prove

Theorem 3.1. *$(T_{Hyp(\tau)};\circ,T_{\sigma_{id}})$ is a monoid which is isomorphic to the monoid of all hypersubstitutions of type τ.*

Proof. We define a mapping $\varphi : Hyp(\tau) \to \mathrm{T}_{Hyp(\tau)}$ by $\sigma \mapsto T_\sigma$. Clearly, φ is well-defined and surjective.

(i): We show that $T_{\sigma_1} \circ T_{\sigma_2} = T_{\sigma_1 \circ_h \sigma_2}$, i.e. $\varphi(\sigma_1 \circ_h \sigma_2) = \varphi(\sigma_1) \circ \varphi(\sigma_2)$. Indeed, we have

$$\begin{aligned}(t,t") \in T_{\sigma_1} \circ T_{\sigma_2} &\Leftrightarrow \exists t'((t,t') \in T_{\sigma_2} \text{ and } (t',t") \in T_{\sigma_1})\\ &\Leftrightarrow t'=\hat{\sigma}_2[t] \text{ and } t"=\hat{\sigma}_1[t'] \Leftrightarrow t" = \hat{\sigma}_1[\hat{\sigma}_2[t]]\\ &\Leftrightarrow t" = (\sigma_1 \circ_h \sigma_2)\hat{}[t] \Leftrightarrow (t,t") \in T_{\sigma_1 \circ_h \sigma_2}.\end{aligned}$$

This shows that $T_{Hyp(\tau)}$ is closed under composition and that φ preserves the operation.

(ii): We show that φ is one-to-one.
Assume that $T_{\sigma_1} = T_{\sigma_2}$. Then for all $t \in W_\tau(X)$ we have $\hat{\sigma}_1[t] = \hat{\sigma}_2[t]$. But this means that for all operation symbols f_i we also have

$$\hat{\sigma}_1[f_i(x_i,\dots,x_{n_i})] = \sigma_1(f_i) = \sigma_2(f_i) = \hat{\sigma}_2[f_i(x_i,\dots,x_{n_i})]$$

and, therefore, $\sigma_1 = \sigma_2$.

Since $T_{\sigma_1} \circ T_{\sigma_2} = T_{\sigma_1 \circ_h \sigma_2}$, the tree transformation $T_{\sigma_{id}}$ is an identity element with respect to the composition $\circ$. ∎

Theorem 3.1 allows us to describe properties of the relation T_σ by properties of the hypersubstitution σ and conversely.

Theorem 3.2. *Let $\sigma \in Hyp(\tau)$ be a hypersubstitution of type τ and let T_σ be the corresponding tree transformation. Then*

(i) *T_σ is transitive iff σ is idempotent,*

(ii) *T_σ is reflexive iff $\sigma = \sigma_{id}$,*

(iii) *T_σ is symmetric iff $\sigma \circ_h \sigma = \sigma_{id}$.*

The proofs are straightforward and left to the reader. ∎

Remark that these propositions are true for any function φ and the corresponding relation T_φ.

A tree transformation is called *injective* if σ is injective, i.e., if from $\hat{\sigma}[t] = \hat{\sigma}[t']$ follows $t = t'$, and T_σ is called *surjective* if σ is surjective.

In general, the range of a tree transformation σ, i.e. the set

$$\hat{\sigma}(W_\tau(X)) = \{t' \mid \exists t \in W_\tau(X)(t' = \hat{\sigma}[t])\},$$

is a subset of $W_\tau(X)$. Therefore, we consider T_σ as a relation between $W_\tau(X)$ and $\hat{\sigma}(W_\tau(X))$, i.e. $T_\sigma = W_\tau(X) \times \hat{\sigma}(W_\tau(X))$, to get surjectivity. We notice that $T_\sigma \circ (T_\sigma)^{-1} = T_{\sigma_{id}} = \triangle_{W_\tau(X)}$ and that $(T_\sigma)^{-1} \circ T_\sigma = \{(t, t') \mid \hat{\sigma}[t] = \hat{\sigma}[t']\} = ker\sigma$ (the kernel of σ). Then we have:

Let $\sigma \in Hyp(\tau)$ be a hypersubstitution of type τ, consider a tree transformation of the form $T_\sigma = W_\tau(X) \times \hat{\sigma}(W_\tau(X))$. Then T_σ is bijective iff $ker\sigma = \triangle_{W_\tau(X)} = T_{\sigma_{id}}$.

In [2] the authors characterized for the type $\tau = (n)$, $n \geq 2$, all hypersubstitutions for which $ker\sigma$ is equal to $\triangle_{W_\tau(X)}$.

It turns out that for $n \geq 2$ this is the case iff $var\sigma(f) = \{x_1, \ldots, x_n\}$, where $var\sigma(f)$ is the set of all variables occurring in $\sigma(f)$. Such hypersustitutions are called *regular.* Let $Reg(\tau), \tau = (n)$, where $n \geq 2$, be the set of all regular hypersubstitutions of type $\tau = (n)$. It is easy to see that $Reg(\tau)$ forms a submonoid of $Hyp(\tau)$ and then, by Theorem 3.1, $T_{Reg(\tau)}$ is a submonoid of $T_{Hyp(\tau)}$ and we have

Corollary 3.3. *Let σ be a hypersubstitution of type $\tau = (n)$, for $n \geq 2$, and let T_σ be the tree transformation defined by σ. Then T_σ is bijective iff σ is regular, i.e. $\sigma \in Reg(\tau)$, and therefore $T_\sigma \in T_{Reg(\tau)}$.* ■

For types different from $\tau = (n)$ there is no explicit description of all hypersubstitutions with $ker\sigma = \Delta_{W_\tau(X)}$, but in [4] one can find an algorithmic solution of this problem.

We remark that $ker\sigma = (T_\sigma)^{-1} \circ T_\sigma$ is always a fully invariant congruence relation on the absolutely free algebra $\mathcal{F}_\tau(X)$ of type τ, i.e. an equational theory.

4. TREE TRANSFORMATIONS OF TYPE $\tau = (2)$

In this section, we want to consider tree transformations defined by hypersubstitutions of type $\tau = (2)$. Let f be the binary operation symbol and let $X_2 = \{x_1, x_2\}$ be a two-element alphabet of variables. Here σ_t for a term

$t \in W_{(2)}(X)$ means that the hypersubstitution σ maps the binary operation symbol f to the term t. In [3] the semigroup properties of the monoid $Hyp(2)$ of all hypersubstitutions of type $\tau = (2)$ were studied.

Let $W_{(2)}(\{x_i\})$ be the set of all terms built up by using only the variable $x_i, i = 1, 2$. Then we define the following sets of hypersubstitutions of type $\tau = (2)$:

$E_{x_1} := \{\sigma \mid \sigma : f \mapsto f(x_1, u) \text{ and } u \in W_{(2)}(\{x_1\})\}$,

$E_{x_1} := \{\sigma \mid \sigma : f \mapsto f(v, x_2) \text{ and } u \in W_{(2)}(\{x_2\})\}$,

$E = E_{x_2} \cup E_{x_2}$,

$M = \{\sigma_{x_1}, \sigma_{x_2}, \sigma_{id}, \sigma_{f(x_2,x_1)}, \sigma_{f(x_1,x_1)}, \sigma_{f(x_2,x_2)}\}$.

A hypersubstitution σ has infinite order if for any $n \geq 1, \sigma^n \neq \sigma^{n+1}$ Then in [3] the following was proved:

Proposition 4.1 ([3]). *Let $\sigma \in Hyp(2)$. Then*

(i) *σ is idempotent iff $\sigma \in E \cup \{\sigma_{x_1}, \sigma_{x_2}, \sigma_{id}\}$;*

(ii) *σ has infinite order iff $\sigma \in Hyp(\tau) \setminus (E \cup M)$.* ∎

As a consequence we have

Corollary 4.2. *Let $\sigma \in Hyp(2)$. Then*

(i) *T_σ is transitive iff $\sigma \in E \cup \{\sigma_{x_1}, \sigma_{x_2}, \sigma_{id}\}$;*

(ii) *T_σ is symmetric iff $\sigma = \sigma_{x_2x_1}$;*

(iii) *T_σ has infinite order iff $\sigma \in Hyp(\tau) \setminus (E \cup M)$.* ∎

Moreover, (i), (ii), (iii) describe all possible cases for T_σ if $\sigma \in Hyp(2)$. If $\sigma \neq \sigma_{id}$, then T_σ is either transitive or symmetric, or has infinite order.

We remark that in [3] some properties of the monoid $Hyp(2)$ concerning Green's relations were proved which can also be applied to the monoid $T_{Hyp(2)}$ of all tree transformations of type $\tau = (2)$.

REFERENCES

[1] K. Denecke, J. Koppitz and St. Niwczyk, *Equational Theories generated by Hypersubstitutions of Type* (n), Internat. J. Algebra Comput., to appear.

[2] K. Denecke, D. Lau, R. Pöschel and D. Schweigert, *Hyperidentities, Hyperequational classes and clone congruences*, Contributions to General Algebra **7** (1991), 97–118.

[3] K. Denecke and S.L. Wismath, *The monoid of hypersubstitutions of type (2)*, Contributions to General Algebra **10** (1998), 109–126.

[4] K. Denecke and S.L. Wismath, *Hyperidentities and Clones*, Gordon and Breach Sci. Publ., Amsterdam 2000.

[5] F. Gécseg and M. Steinby, *Tree Automata,* Akadémiai Kiadó, Budapest 1984.

[6] J.W. Thatcher, *Tree Automata: an informal survey*, p. 143–172 in: "*Currents in the theory of computing*", Prentice-Hall, Englewood Cliffs, NJ, 1973.

Received 4 June 2001
Revised 30 November 2001