Department of Foundations Computer Science

RESEARCH RESULTS ON TREE PROCESSING DEVICES 2000–2003

Zoltán Fülöp

I. TREE TRANSDUCERS

In [1] we prove that the domains of partial attributed tree transducers are the tree languages recognized by tree walking automata in universal acceptance mode. Moreover, we prove that the domains of partial attributed tree transducers having monadic output are the tree languages recognized by deterministic tree walking automata.

In [3] we consider attributed tree transducers of type simple multi-visit, multi-sweep, multi-alternating pass, and multi-pass following the definition of the same subclasses of attribute grammars. We characterize the tree transformation classes induced by the considered four subclasses of attributed tree transducers in terms of macro tree transducers. Namely, we define simple multi-visit, multi-sweep, multi-alternating pass and multi-pass macro tree transducers and prove formally that the attributed tree transducers and the macro tree transducers of the same type induce the same tree transformation class. Also we give an inclusion diagram of the tree transformation classes induced by the above and some further fundamental types of attributed and of macro tree transducers.

In [5] we consider the closure $UCI(Rel)$ of the class of relabelling tree transformations, under U=union, C=composition and I=iteration. We give a characterization of $UCI(Rel)$ in terms of a short expression built up from Rel with composition and iteration. We also give a characterization of $UCI(Rel)$ in terms of one-step rewrite relations of very simple term rewrite systems. We give a similar characterization of $UC(FRel_+)$ where $FRel_+$ is the class consisting of the transitive closures of all functional relabelling tree transformations. Finally we show that $UCI(Rel) = UCI(FRel)$.

In [7] we deal with shape-preserving top-down tree transducers. As top-down tree transducers generalize GSM's, shape preserving top-down tree transducers naturally generalize length preserving GSM's. For instance, top-down finite state relabelling tree transducers are shape preserving top-down tree transducers. We show that a top-down tree transducer is shape preserving if and only if it is equivalent to a top-down finite state relabelling tree transducer. We also prove that it is decidable if a top-down tree transducer is shape preserving.

II. TREE SERIES TRANSDUCERS

In [6] we generalize bottom-up tree transducers and top-down tree transducers to the concept of *bottom-up tree series transducer* and *top-down tree series transducer*, respectively, by allowing formal tree series as output rather than trees, where a formal tree series is a mapping from output trees to some semiring. We associate two semantics with a tree series transducer: a mapping which transforms trees into tree series (for short: tree to tree series transformation or t-ts transformation), and a mapping which transforms tree series into tree series (for short: tree series transformation or ts-ts transformation).

We show that the standard case of tree transducers is reobtained by choosing the boolean semiring under the t-ts semantics. Moreover, we show that certain fundamental constructions and results concerning bottom-up and top-down tree transducers can be generalized for the corresponding tree series transducers. Among others, we prove that polynomial bottom-up t-ts transformations can be characterized by the composition of finite state relabelling t-ts transformations and boolean homomorphism t-ts transformations. Moreover, we prove that every deterministic top-down t-ts transformation can be characterized by the composition of a boolean homomorphism t-ts transformation and a deterministic linear top-down t-ts transformation. We prove that deterministic top-down t-ts transformations are closed under right composition with nondeleting and linear deterministic top-down t-ts transformations and are closed under left composition with boolean and total deterministic top-down t-ts transformations. Finally we show that nondeleting linear bottom-up and nondeleting linear top-down tree series transducers generate the same t-ts transformation class.

In [8] we consider tree series transducers which were introduced in [6], and define the tree-to-tree series transformations computed by them in two different ways. One of the definitions is based on the $\leftarrow$-substitution of tree series taken from [6] while the other one is based on a new tree series substitution introduced in this paper. This new substitution is called $\overset{o}{\leftarrow}$-substitution and the main difference between the $\leftarrow$- and the $\overset{o}{\leftarrow}$-substitutions is that the first one does not take into account the number of the occurrences of the substitution variables while the second one does. We compare the two different ways of computing tree-to-tree series transformations and show that, for the $\overset{o}{\leftarrow}$-substitution, fundamental relations from the theory of tree transducers carry over to tree series transducers.

In [9] we study bottom-up and top-down tree series transducers over a semiring A and denote the tree series transformation classes computed by them by $BOT_{t-ts}(A)$ and $TOP_{t-ts}(A)$, respectively. We present the inclusion diagram of the classes p-$BOT^n_{t-ts}(A)$, p-$TOP^n_{t-ts}(A)$, p-$BOT^{n+1}_{t-ts}(A)$, and p-$TOP^{n+1}_{t-ts}(A)$ and prove its correctness, where A is a commutative izz-semiring (izz = idempotent, zero-divisor free, and zero-sum free) and the prefix p stands for polynomial. This inclusion diagram implies the properness of the following four hierarchies:

$$p\text{-}TOP_{t-ts}(A) \subseteq p\text{-}TOP^2_{t-ts}(A) \subseteq p\text{-}TOP^3_{t-ts}(A) \subseteq \dots$$

$$p\text{-}BOT_{t-ts}(A) \subseteq p\text{-}BOT^2_{t-ts}(A) \subseteq p\text{-}BOT^3_{t-ts}(A) \subseteq \dots$$

$$p\text{-}TOP_{t-ts}(A) \subseteq p\text{-}BOT^2_{t-ts}(A) \subseteq p\text{-}TOP^3_{t-ts}(A) \subseteq p\text{-}BOT^4_{t-ts}(A) \subseteq \ldots$$

$$p\text{-}BOT_{t-ts}(A) \subseteq p\text{-}TOP^2_{t-ts}(A) \subseteq p\text{-}BOT^3_{t-ts}(A) \subseteq p\text{-}TOP^4_{t-ts}(A) \subseteq \ldots,$$

where the first hierarchy generalizes the famous top-down tree transformation hierarchy of J. Engelfriet. As the second main result we prove that the first two hierarchies are proper even for arbitrary (i.e., not necessarily commutative) izz-semirings.

In [6, 8] the semantics of tree series transducers was defined in an algebraic framework, more precisely, as an initial algebra semantics. In [10] we suggest an alternative approach by introducing weighted tree transducers, of which the semantics is defined in an operational style. A weighted tree transducer is a tree transducer each (term rewriting) rule of which is associated with a weight taken from a semiring. Along a successful derivation the weights of the involved rules are multiplied and, for every pair of input tree and output tree, the weights of its successful derivations are summed up. We show in a constructive way that the two approaches, i.e., tree series transducers and weighted tree transducers, are semantically equivalent for both, the top-down and the bottom-up case.

III. TREE AUTOMATA AND GROUND TERM REWRITING

In [4] we show that the component hierarchy of chain-free distributed regular tree grammars cooperating with terminal strategy is infinite with respect to tree language generating capacity. More exactly, we prove that $cf\text{-}CD\text{-}RTG_{\dashv}(n) \subset cf\text{-}CD\text{-}RTG_{\dashv}(2(n-1)^2+3)$, where $n \geq 1$ and $cf\text{-}CD\text{-}RTG_{\dashv}(n)$ denotes the class of tree languages generated by chain-free distributed regular tree grammars of at most n components cooperating with terminal strategy.

In [2] we consider restricted versions of ground tree transducers: total, deterministic, and symmetric subclasses and all other subclasses created by applying any combination of these restrictions. We present the inclusion diagram of the tree transformation classes induced by these restricted ground tree transducers.

We show that the following four classes of term relations are the same: (i) tree transformations induced by symmetric deterministic ground tree transducers, (ii) congruence relations on term algebras induced by reduced ground term rewriting systems, (iii) congruence relations on term algebras induced by convergent ground term rewriting systems, and (iv) finitely generated congruence relations on term algebras.

As a by-product of our results, we obtain a new ground completion algorithm.

Moreover, we show that the following three classes of term relations on term algebras with at least one non-nullary function symbol are also the same: (i) tree transformations induced by total symmetric deterministic ground tree transducers, (ii) congruence relations on term algebras of finite index, (iii) finitely generated congruence relations on term algebras of which the trunk is the whole set of terms.

CONFERENCE ORGANIZATION AND EDITORIAL WORK

In the period 2000-2003 Z. Fülöp served as the member of the Program Committee of the conferences Mathematical Foundations of Computer Science 2000, European Conference on Intelligent Systems and Technologies 2002, Developments in Language Theory 2002 and 2003 (co-chair), Automata and Formal Languages 2002. He was the coeditor of the proceedings [11] of DLT 03. He has been the editor of the journals Acta Cybernetica, Informatica, and Grammas.

ACKNOWLEDGEMENT

The above research has been supported by several grants from the National Foundation of Hungary for Scientific Research, the Exchange Program of the DAAD and the Hungarian Scholarship Committee and the International Quality Network program of the DAAD.

REFERENCES

[1] Z. Fülöp and S. Maneth, Domains of Partial Attributed Tree Transducers (S. Manethval), *Inf. Process. Letters*, **73** (2000) 175-180.

[2] Z. Fülöp and S. Vágvölgyi, Restricted Ground Tree Transducers, *Theoret. Comput. Science*, **250** (2001) 219-233.

[3] Z. Fülöp, Characterizing Important Subclasses of Attributed Tree Transformations by Macro Tree Transducers, *Journal of Automata, Languages and Combinatorics* **6** (2001) 145-189.

[4] G. Dányi and Z. Fülöp, The Component Hierarchy of Chain-Free Cooperating Distributed Regular Tree Grammar Systems, *Theoret. Comput. Science*, **262** (2001) 229-240.

[5] Z. Fülöp and A. Terlutte, Iterated Relabeling Tree Transducers, *Theoret. Comput. Science*, 276 (2002) 221-244.

[6] J. Engelfriet, Z. Fülöp and H. Vogler, Bottom-up and Top-down Tree Series Transformations,*Journal of Automata, Languages and Combinatorics*, **7** (2002) 11-70.

[7] Z. Fülöp and Zs. Gazdag, Shape Preserving Top-Down Tree Transducers, Theoret. Comput. Science, **304** (2003) 315-339.

[8] Z. Fülöp and H. Vogler, Tree Series Transformations that Respect Copying, Theory of Computing Systems, **36** (2003) 247-293.

[9] Z. Fülöp, Zs. Gazdag and H. Vogler, Hierarchies of Tree Series Transformations, *Theoret. Comput. Science*, accepted for publication.

[10] Z. Fülöp and H. Vogler, Weighted Tree Transducers (with H. Vogler), *Journal of Automata, Languages and Combinatorics*, accepted for publication.

[11] Proceedings of DLT 03 (editor Z. Ésik and Z. Fülöp), Lecture Notes in Computer Science 2710, Springer-Verlag, 2003.