TREE AUTOMATA AND TERM REWRITE SYSTEMS

Sándor Vágvölgyi

I. TERM REWRITE SYSTEMS EFFECTIVELY PRESERVING RECOGNIZABILITY

Tree automata and recognizable tree languages proved to be an efficient tool in the theory of term rewrite systems, see [11] for an overview. Let Σ be a ranked alphabet, let R be a term rewrite system over Σ, and let L be a tree language over Σ. Then $R^*_\Sigma(L)$ is the set of descendants of trees in L. When Σ is apparent from the context, we simply write $R^*(L)$ rather than $R^*_\Sigma(L)$. A term rewrite system R over Σ preserves Σ-recognizability, if for each recognizable tree language L over Σ, $R^*(L)$ is recognizable. The signature $sign(R)$ of a term rewrite system R is the ranked alphabet consisting of all symbols appearing in the rules of R. In [10] Gilleron showed that for a term rewrite system R it is not decidable if R preserves $sign(R)$-recognizability. Otto [16] showed that it is not decidable for a term rewrite system R whether R preserves recognizability. A term rewrite system R over Δ preserves recognizability, if for each ranked alphabet Σ with $sign(R) \subseteq \Sigma$, R preserves Σ-recognizability. We [12] showed that there is a ranked alphabet Σ and there is a linear term rewrite system R over Σ such that R preserves Σ-recognizability but does not preserve recognizability.

Let R be a term rewrite system over a ranked alphabet Σ. We say that R effectively preserves Σ-recognizability if for a given tree automaton $\mathcal{B}$ over Σ, we can effectively construct a tree automaton $\mathcal{C}$ over Σ such that $L(\mathcal{C}) = R^*_\Sigma(L(\mathcal{B}))$. Let R be a term rewrite system over a ranked alphabet Δ. We say that R effectively preserves recognizability if for any given ranked alphabet Σ with $sign(R) \subseteq \Sigma$ and any given tree automaton $\mathcal{B}$ over Σ, we can effectively construct a tree automaton $\mathcal{C}$ over Σ such that $L(\mathcal{C}) = R^*_\Sigma(L(\mathcal{B}))$.

In spite of the undecidability results of Gilleron and Otto, we know several term rewrite systems which preserve recognizability. Brainerd [2] showed that ground term rewrite systems over any ranked alphabet Σ effectively preserve Σ-recognizability. Gallier and Book [7] introduced the notion of a monadic term rewrite system, and Salomaa [19] showed that linear monadic term rewrite systems over any ranked alphabet Σ effectively preserve Σ-recognizability. A term rewrite system is monadic if each left-hand side is of depth at least 1 and each right-hand side is of depth at most 1. Coquidé et al [3] defined the concept of a semi-monadic term rewrite system generalizing the notion of a monadic rewrite system and the notion of a ground term rewrite system. A term rewrite system R over Σ is semi-monadic if, for every rule $l \to r$ in R, $depth(l) \geq 1$ and either $depth(r) = 0$ or $r = f(y_1, \ldots, y_k)$, where $f \in \Sigma_k$, $k \geq 1$, and for each $i \in \{1, \ldots, k\}$, either y_i is a variable (i.e., $y_i \in X$) or y_i is a ground term (i.e., $y_i \in T_\Sigma$). It is immediate that each monadic term rewrite system is semi-monadic as well. Coquidé et al [3] showed that linear semi-monadic term rewrite systems over any ranked alphabet Σ effectively preserve Σ-recognizability. We [12] generalized even further the concept of a semi-monadic term rewrite system introducing the concept of a generalized semi-monadic term rewrite system (gsm term rewrite system for short). We [12] showed that a linear gsm (lgsm) term rewrite system R over Δ effectively preserves recognizability.

Let R be a term rewrite system over $sign(R)$, and let $\Sigma = \{ f, \sharp \} \cup sign(R)$, where $f \in \Sigma_2 - sign(R)$ and $\sharp \in \Sigma_0 - sign(R)$. We [12] showed that R effectively preserves Σ-recognizability if and only if R effectively preserves recognizability. We [13] improved this result for left-linear term rewrite systems. We showed the following. Let R be a left-linear term rewrite system over $sign(R)$, and let $\Sigma = \{ g, \sharp \} \cup sign(R)$, where $g \in \Sigma_1 - sign(R)$ and $\sharp \in \Sigma_0 - sign(R)$. Then R effectively preserves Σ-recognizability if and only if R effectively preserves recognizability. This result makes it easier to show that a given left-linear term rewrite system effectively preserves recognizability. Furthermore, R preserves Σ-recognizability if and only if R preserves recognizability.

We [26] introduced and studied the half-monadic term rewrite system, which is an extension of the right-ground term rewrite system and is a slight extension of the semi-monadic term rewrite system, see [3]. Takai, Kaji, and Seki [21] introduced the concept of the finite path overlapping term rewrite system, and showed that right-linear finite path overlapping term rewrite systems effectively preserve recognizability. A half-monadic term rewrite system is a finite path overlapping term rewrite system as well. Hence right-linear half-monadic term rewrite systems effectively preserve recognizability. We [26] showed that termination and convergence are decidable properties for right-linear half-monadic term rewrite systems.

Let R be a right-linear half-monadic term rewrite system over a ranked alphabet Σ such that for every rule $l \to r$ in R, l is not a variable. We introduced two weight functions. Let $p \in T_\Sigma$ be a minimal height ground term starting an infinite reduction sequence. We introduced the weight function $\| \|_{p,R} : R^*(\{p\}) \to N$. For any $t \in R^*(\{p\})$, $\|t\|_{p,R}$ is the length of the longest reduction sequence starting from t, where we do not rewrite at the root and we do not count the rewrite steps of those immediate subtrees of the root which are subtrees of the descendants of the immediate ground subterms of the right-hand sides of the rules in R. We showed that $\to_R^*$ does not increase $\| \|_{p,R}$. We gave a condition which implies that $\to_R^*$ decreases $\| \|_{p,R}$. For an arbitrary $p \in T_\Sigma$, we introduced the weight function $\langle \rangle_{p,R} : R^*(\{p\}) \to N$. For any $t \in R^*(\{p\})$, $\langle t\rangle_{p,R}$ is the number of occurrences of those subtrees which are not subtrees of the descendants of the immediate ground subterms of the right-hand sides of the rules in R. We showed that $\to_R^*$ does not increase $\langle \rangle_{p,R}$. We gave a condition which implies that $\to_R^*$ decreases $\langle \rangle_{p,R}$.

Using the effective preserving recognizability property of R, we show that for a ground term $t \in T_\Sigma$ it is decidable if $t \in T_\Sigma$ starts an infinite reduction sequence. Using weight functions $\| \|_{p,R}$ and $\langle \rangle_{p,R}$, we constructed a finite set $W \subseteq T_\Sigma$ of terms such that R is terminating if and only if no term in W starts an infinite reduction sequence. We constructed W in two steps. First we put in W the subtrees of the descendants of the ground subterms of the right-hand sides of the rules in R. Then we put in W all ground terms obtained from the left-hand sides of the rules in R by replacing the variables with nullary symbol \$ and subtrees of the descendants of the ground subterms of the right-hand sides of the rules in R. In this way we decide if R is terminating. Hence we get that termination is a decidable property for right-linear half-monadic term rewrite systems. As R preserves recognizability we can decide if R is locally confluent. Hence convergence is a decidable property for right-linear half-monadic term rewrite systems.

II. GROUND TERM REWRITE SYSTEMS

In paper [22], we showed that a reduced ground term rewrite system over some ranked alphabet Σ has a good property: if we omit any rules, the omitted rules and the remaining rules generate congruence relations such that their intersection is the identity relation on T_Σ. In general, given two ground term rewrite systems A and B over some ranked alphabet Σ with $\leftrightarrow_A^* \subseteq \leftrightarrow_B^*$, one may want to eliminate all nontrivial pairs of $\leftrightarrow_A^*$ from $\leftrightarrow_B^*$. In other words, one may want to construct a ground term rewrite system (gtrs for short) C over Σ such that $\leftrightarrow_{A\cup C}^* = \leftrightarrow_B^*$ and that $\leftrightarrow_A^* \cap \leftrightarrow_C^*$ is the identity relation on T_Σ. Let A and B be gtrs's over some ranked alphabet Σ with $\leftrightarrow_A^* \subseteq \leftrightarrow_B^*$. We say that a gtrs C over Σ is a congruential complement of A for B, if $\leftrightarrow_{A\cup C}^* = \leftrightarrow_B^*$ and $\leftrightarrow_A^* \cap \leftrightarrow_C^*$ is the identity relation over T_Σ. We note that A may have more than one congruential complement for B, which are pairwise nonequivalent. Our [22] main result was the following. Given gtrs's A, B, C over some ranked alphabet Σ with $\leftrightarrow_A^* \subseteq \leftrightarrow_B^*$, one can effectively decide if C is a congruential complement of A for B.

Dauchet et al [4] have introduced the notion of the ground tree transducer as a pair (A, B) of tree automata. The importance of ground tree transducers is in that they can simulate ground term rewriting: in [4] it was shown that for each ground term rewriting system R over a ranked alphabet Σ, one can effectively construct a ground tree transducer (A, B) over Σ such that $\to_R^*$ is equal to the tree transformation $\tau_{(A,B)}$ induced by (A, B). The authors of [4] have used this result to show that the confluence property of ground term rewrite systems is decidable, cf. [17]. From now on we write $\pi(A)$ for $\tau_{(A,A)}$.

In paper [9] we considered restricted versions of ground tree transducers. Our motivation was that studying restricted versions of a class of machines frequently gives a deeper insight into the working of the unrestricted class. As usual for a class of tree transducers, we considered the total and the deterministic subclasses. We call a ground tree transducer (A, B) deterministic (total) if the tree automata A and B are deterministic (total). We also considered symmetric ground tree transducers, which are of the form (A, A). Moreover, we considered the eight classes of ground tree transducers obtained by combining these three properties in all possible ways.

As the first result, we compared the expressive power of the eight classes by presenting the full inclusion diagram of the tree transformation classes induced by them. Then we [9] showed that the following four classes of term relations are the same: (i) the tree transformations induced by symmetric deterministic ground tree transducers, (ii) the congruence relations on term algebras induced by reduced ground term rewriting systems, (iii) the congruence relations on term algebras induced by convergent ground term rewriting systems, and (iv) the finitely generated congruence relations on term algebras.

In [6], Fülöp and Vágvölgyi showed the following. For every ground term equation system E over a ranked alphabet Σ, which is just a finite binary relation on T_Σ, one can effectively construct a deterministic tree automaton A over Σ such that $\leftrightarrow_E^*$, i.e., the congruence relation induced by E on the Σ-term algebra, is equal to the tree transformation $\pi(A)$ induced by the symmetric deterministic ground tree transducer (A, A).

In the proof of the inclusion (i) $\subseteq$ (ii) we constructed, for a given deterministic tree automaton A over a ranked alphabet Σ, a reduced ground term rewriting system R over Σ such that the congruence relation $\leftrightarrow_R^*$ generated by R is equal to the tree transformation $\pi(A)$.

Thus, as a by-product of our results, we obtained a new ground completion algorithm. Given ground term equation system E, we constructed a reduced ground term rewriting system equivalent to E in two steps. In the first step, presented in [6], we computed a symmetric deterministic ground tree transducer (A, A) such that $\pi(A) = \leftrightarrow_E^*$. Then we constructed as in the proof of (i) $\subseteq$ (ii) a reduced ground term rewriting system R such that $\pi(A) = \leftrightarrow_R^*$. Hence $\leftrightarrow_E^* = \leftrightarrow_R^*$. This ground completion parallels to Snyder's fast algorithm, see [20] and the results of the papers [18] and [14].

It can easily be proved that a binary relation τ over T_Σ can be induced by a total symmetric deterministic ground tree transducer if and only if τ is a congruence of finite index on the Σ-term algebra. In addition, we [9] showed that any finitely generated congruence relation τ over the Σ-term algebra is of finite index if $trunk(\tau) = T_\Sigma$. The concept of the trunk was introduced in [8] and proved to be useful in studying congruence relations over terms. Furthermore, we proved that for a term algebra with at least one non-nullary function symbol the congruence relations of finite index are exactly the finitely generated congruence relations with trunks equal to the set of all terms.

Fülöp and Vágvölgyi [9] constructed, for a given deterministic tree automaton A over Σ, a reduced ground term rewrite system R over Σ such that $\leftrightarrow_R^*$ is equal to the ground tree transformation $\pi(A)$. They also showed that the class of ground tree transformations induced by tree automata properly contains the class of ground tree transformations induced by deterministic tree automata.

A tree automaton A over Σ induces another binary relation, $\theta(A) = \leftrightarrow_A^* \cap T_\Sigma \times T_\Sigma$ over T_Σ. It is the restriction of the congruence $\leftrightarrow_A^*$ on the $\Sigma \cup A$-term algebra to terms over Σ. Hence $\theta(A)$ is a congruence on the Σ-term algebra. Fülöp and Vágvölgyi [5] showed the following. For every ground term rewrite system R over Σ, one can effectively construct a deterministic tree automaton A over Σ such that $\leftrightarrow_R^* = \theta(A) = \pi(A)$. We [25] showed that for any deterministic tree automaton A over Σ, $\theta(A) = \pi(A)$.

For a tree automaton A over Σ, we [25] defined a congruence relation on A, called the determiner of A. We showed that we can effectively compute the determiner ρ, and construct the quotient tree automaton A/ρ. We called A/ρ the principal quotient of A, and showed that A/ρ is a deterministic tree automaton. We [25] showed the following results.

- For any tree automata A and B over Σ and homomorphism $\phi : A \to B$, $\theta(A) \subseteq \theta(B)$ and $\pi(A) \subseteq \pi(B)$.

- For any tree automaton A, the following two statements are equivalent.

  (i) $\pi(A) = \pi(A/\rho)$.

  (ii) There is a deterministic tree automaton B such that $\pi(A) = \pi(B)$.

- For any connected tree automaton A, Conditions (i) and (ii) are equivalent to Condition (iii):

  (iii) $\theta(A) = \pi(A)$.

- For any tree automaton A over Σ, Conditions (i), (ii), and (iii) are decidable.

We [23] showed that it is decidable for any given ground term rewrite systems R and S if there is a ground term rewrite system U such that $\leftrightarrow_U^* = \leftrightarrow_R^* \cap \leftrightarrow_S^*$. If the answer is yes, then we can effectively construct such a ground term rewrite system U. In other words, for any given finitely generated congruences ρ and τ over the term algebra, it is decidable if $\rho \cap \tau$ is a finitely generated congruence. If the answer is yes, then we can effectively construct a ground term rewrite system U such that $\leftrightarrow_U^* = \rho \cap \tau$.

In [24] we studied rewriting of ground terms. For an arbitrary term rewrite system R over a ranked alphabet Σ, we restricted the rewriting relation $\to_R$ to ground terms. We introduced the relation $\to_{R,g} = \to_R \cap (T_\Sigma \times T_\Sigma)$. We showed that $\leftrightarrow_{R,g}^* = \leftrightarrow_R^* \cap (T_\Sigma \times T_\Sigma)$. We showed that for a given term rewrite system R and a given ground term rewrite system S over a ranked alphabet Σ it is decidable if $\leftrightarrow_{R,g}^* \subseteq \leftrightarrow_S^*$. We showed that for a given left-linear right-ground term rewrite system R over a ranked alphabet Σ it is decidable if there is a ground term rewrite system S over Σ such that $\to_{R,g}^* = \to_S^*$. If the answer is yes, then one can effectively construct a gtrs S over Σ such that $\to_{R,g}^* = \to_S^*$.

In paper [1] we studied that version of the ground tree transducer game where the same tree automaton appears as the first and second component of the associated ground tree transducer. We gave conditions which imply that Beta has a winning strategy. Furthermore, we showed the following decidability result. Given a ground tree transducer game where the underlying tree automaton A cannot evaluate some tree into a state or A is deterministic, we can decide which player has a winning strategy. Moreover, whatever player has a winning strategy, we can effectively construct a partial recursive winning strategy for him.

ACKNOWLEDGEMENTS

The mentioned research tasks were supported by the following grants: FKFP 0095/1997 of the Hungarian Cultural and Educational Ministry, OTKA T022994, OTKA T030084 of the Research Foundation of Hungary, the Bolyai János Research Scholarship of the Hungarian Academy of Sciences, and the Széchenyi István Scholarship of the Hungarian Ministry of Education.

REFERENCES

[1] J. Apró and S. Vágvölgyi, The ground tree transducer game with identical tree automata, *Fundamenta Informaticae*, **45** (2001) 157-172.

[2] W.S. Brainerd, Tree generating regular systems, *Information and Control* **14** (1969) 217-231.

[3] J.L. Coquidé, M. Dauchet, R. Gilleron, and S. Vágvölgyi, Bottom-up tree pushdown automata: classification and connection with term rewrite systems, *Theoretical Computer Science* **127** (1994) 69-98.

[4] M. Dauchet, P. Heuillard, P. Lescanne and S. Tison, Decidability of the confluence of finite ground term term rewrite systems and of other related term rewrite systems, *Information and Control* **88** (1990) 187-201.

[5] Z. Fülöp and S. Vágvölgyi, Congruential tree languages are the same as recognizable tree languages - A proof for a theorem of D. Kozen, *Bulletin of the EATCS*, **39** (1989) 175-185.

[6] Z. Fülöp and S. Vágvölgyi, Ground term rewriting rules for the word problem of ground term equations, *Bulletin of the EATCS*, **45** (1991) 186-201.

[7] J.H. Gallier and R.V. Book, Reductions in tree replacement systems *Theoretical Computer Science* **37** (1985) 123-150.

[8] Z. Fülöp and S. Vágvölgyi, Minimal Equational Representations of Recognizable Tree Languages, *Acta Informatica*, **34** (1997) 59-84.

[9] Z. Fülöp and S. Vágvölgyi, Restricted Ground Tree Transducers, *Theoretical Computer Science*, **250** (2001) 219-233.

[10] R. Gilleron, Decision problems for term rewriting systems and recognizable tree languages, Proc. STACS'91, *Lecture Notes in Computer Science* **480**, (1991) 148-159.

[11] R. Gilleron and S. Tison, Regular tree languages and term rewrite systems, *Fundamenta Informaticae* **24** (1995) 157-175.

[12] P. Gyenizse and S. Vágvölgyi, Linear Generalized Semi-monadic Rewrite Systems Effectively Preserve Recognizability, *Theoretical Computer Science*, **194** (1998) 87-122.

[13] P. Gyenizse, S. Vágvölgyi, A property of left-linear rewrite systems preserving recognizability, *Theoretical Computer Science*, **242** (2000) 477-498.

[14] D. Kapur, Shostak's Congruence Closure as Completion, in Rewriting Techniques and Applications: in Proc. of the 8th International Conference, (Sitges, Spain, 1997) *Lecture Notes in Computer Science*, Vol. 1232, Springer Publishing Company, New York, 1997, 23-37.

[15] D. Kozen, Complexity of finitely presented algebras, *Proc. 9th Annual ACM Symp. on Theory of Computing* (Boulder Colorado, 1977) 164-177.

[16] F. Otto, Some undecidability results concerning the property of preserving regularity, *Theoretical Computer Science* **207** (1998) 43-72.

[17] M. Oyamaguchi, The Church-Rosser property for ground term rewriting systems is decidable, *Theoretical Computer Science*, 49 (1987) 43-79.

[18] D. Plaisted and A. Sattler-Klein, Proof lengths for equational completion, *Information and Computation*, **125** (1996) 154-170.

[19] K. Salomaa, Deterministic tree pushdown automata and monadic tree rewriting systems, *Journal of Computer and System Sciences* **37** (1988) 367-394.

[20] W. Snyder, A Fast Algorithm for Generating Reduced Ground Rewriting Systems from a set of Ground Equations, *Journal of Symbolic Computation*, 15 (1993) 415-450.

[21] T. Takai, Y. Kaji, and H. Seki, Right-Linear Finite Path Overlapping Term Rewriting Systems Effectively Preserve Recognizability, Rewriting Techniques and Applications, *Lecture Notes in Computer Science* **1833** (2000) 246-260.

[22] S. Vágvölgyi, Congruential Complements of Ground Term Rewrite Systems, *Theoretical Computer Science*, **238** (2000) 247-274.

[23] S. Vágvölgyi, Intersection of Finitely Generated Congruences over Term Algebra *Theoretical Computer Science*, **300** (2003) 209-234.

[24] S. Vágvölgyi, Term Rewriting Restricted to Ground Terms *Theoretical Computer Science*, **302** (2003) 135-165.

[25] S. Vágvölgyi, On Ground Tree Transformations and Congruences Induced by Tree automata *Theoretical Computer Science*, **304** (2003) 449-459.

[26] S. Vágvölgyi, Right-Linear Half-Monadic Term Rewrite Systems, *Theoretical Computer Science*, **309** (2003) 195-211.