MR2150756 (2006c:68129) 68Q70 (68Q45)
Ésik, Zoltán (H-SZEG-C); **Weil, Pascal** (F-BORD-LB)
Algebraic recognizability of regular tree languages. (English summary)
Theoret. Comput. Sci. **340** (2005), *no. 2,* 291–321.

One of the most important open problems in the theory of tree automata and tree languages is the decidability of first order definable tree languages (over the signature of label predicates and the ascendency relation). Let $\mathrm{FO}[<]$ be this family of first order definable tree languages. In the paper under review the authors claim to have found a characterization of $\mathrm{FO}[<]$ by some algebraic structures, called preclones. This was announced earlier in their conference paper [Z. Ésik and P. Weil, in *FST TCS 2003: Foundations of software technology and theoretical computer science*, 195–207, Lecture Notes in Comput. Sci., 2914, Springer, Berlin, 2003; MR2093649 (2005e:68168)]. The present paper provides the background for its sequel ["Algebraic characterization of logically defined tree languages", in preparation] by the same authors. The authors note, however, that it is not clear at the moment if this algebraic characterization may lead to a decidability algorithm for $\mathrm{FO}[<]$; in other words, the aforementioned problem remains open.

A preclone is a many-sorted algebraic structure whose set of sorts is $\mathbb{N}$, the set of non-negative integers. An element of sort n can be combined with some n elements of arbitrary sorts, say of sorts $m_1,\dots,m_n$, and result in an element of sort $\sum_{i=1}^{i=n} m_i$. The signature of preclones also contains a constant symbol for the identity element of sort 1. A prominent example of a preclone is the set of transformations of a fixed set A; elements of A are elements of the preclone of sort 0, and for $m>0$, m-ary transformations $A^m\to A$ are the elements of sort m. Note that the composition operation acts as if no variable appears more than once in any of these transformations: we cannot have a transformation, for example, of the form $x\mapsto f(x,x)$. This is the main difference of a preclone from a clone (where transformations are freely composed with each other).

For this restriction, the authors consider trees in a rather peculiar formalism: for a ranked alphabet Σ and a leaf alphabet X, which is taken to be $\{x_1,\dots,x_m\}$ in the paper, trees are terms in which any x_i appears exactly once—moreover, in the order $x_1\dots x_m$ from left to right. In the first author's earlier paper [Publ. Math. Debrecen **54** (1999), suppl., 711–762; MR1709922 (2000h:68115)], where clones were studied for characterizing families of tree languages, it was claimed that $\mathrm{FO}[<]$ is closed under those inverse morphisms in which no variable occurs twice. So, in that setting, $\mathrm{FO}[<]$ is not a variety (to be defined by clones). In the present paper, this subfamily of $\mathrm{FO}[<]$ (in which no leaf variable appears twice) is a variety of tree languages definable by preclones.

Syntactic preclones (similarly, syntactic clones) of tree languages have more expressive power than their syntactic algebras, or monoids, since the elements of sort 0 of the syntactic preclone of a tree language constitute its syntactic algebra, and the set of elements of sort 1 is its syntactic monoid. No clear comparison of the present algebraic formalism with the other frameworks (algebras, monoids, etc.) is given yet.

Pseudovarieties of preclones and varieties of tree languages are defined to be special classes closed under certain closure properties, and a variety theorem for these classes is proved in the paper (Theorem 5.1). Also some families of tree languages are characterized by preclones, which include

(1) FO-definable tree languages, studied in [M. A. Benedikt and L. Segoufin, in *STACS 2005*, 327–339, Lecture Notes in Comput. Sci., 3404, Springer, Berlin, 2005; MR2151629 (2006b:03046)],

(2) TL(EF), TL(EX), TL(EF + EX) studied in [M. Bojańczyk and I. Walukiewicz, in *CONCUR 2004—concurrency theory*, 131–145, Lecture Notes in Comput. Sci., 3170, Springer, Berlin, 2004], and

(3) FO[<] (announced without detailed proof in [Z. Ésik and P. Weil, op. cit., 2003]).

Here one wonders if the characterizations (1) and (2) above could have been presented in the clone framework; or if there is a reason for preferring preclones to clones in these cases.

The paper is well written, technically demanding, and rather self-contained; the only place where the reader needs to see the references (which are listed in (1) and (2) above) is SubSection 5.2. Some minor typos I noticed are: In Example 2.7 it seems that true_1 is also needed for generating T_{path}; T_l in line -11 on page 298 should be S_l; and $\Sigma' M_k$ in line -11 on page 305 should be ΣM_k.

Reviewed by *Saeed Salehi*

References

1. J. Almeida, On pseudovarieties, varieties of languages, filters of congruences, pseudoidentities and related topics, Algebra Univ. 27 (1990) 333–350. MR1058478 (91h:08011)
2. A. Arnold, M. Dauchet, Théorie des magmoïdes I and II, RAIRO Theoret. Inform. Appl. 12 (1978), 235–257; 3 (1979) 135–154 (in French). MR0510640 (81d:68091a)
3. M. Benedikt, L. Ségoufin, Regular tree languages definable in FO, in: V. Diekert, B. Durand (Eds.), STACS 2005, Lecture Notes in Computer Science, vol. 3404, Springer, Berlin, pp. 327–339. MR2151629 (2006b:03046)
4. S.L. Bloom, Z. Ésik, Iteration Theories, Springer, Berlin, 1993. MR1295433 (95g:68065)
5. M. Bojańczyk, I. Walukiewicz, Characterizing EF and EX tree logics, in: P. Gardner, N. Yoshida (Eds.), CONCUR 2004, Lecture Notes in Computer Science, Vol. 3170, Springer, Berlin, 2004.
6. J.R. Büchi, Weak second-order arithmetic and finite automata, Z. Math. Logik Grundlagen Math. 6 (1960) 66–92. MR0125010 (23 #A2317)
7. J. Cohen, J.-E. Pin, D. Perrin, On the expressive power of temporal logic, J. Comput. System Sci. 46 (1993) 271–294. MR1228808 (94i:68154)
8. H. Comon, M. Dauchet, R. Gilleron, F. Jacquemard, D. Lugiez, S. Tison, M. Tommasi, Tree Automata Techniques and Applications, available on: http://www.grappa.univ-lille3.fr/tata (release October 2002).
9. B. Courcelle, The monadic second-order logic of graphs. I. Recognizable sets of finite graphs, Inform. and Comput. 85 (1990) 12–75. MR1042649 (91g:05107)
10. B. Courcelle, Basic notions of universal algebra for language theory and graph grammars, Theoret. Comput. Sci. 163 (1996) 1–54. MR1407014 (97k:68098)
11. B. Courcelle, The expression of graph properties and graph transformations in monadic second order logic, in: G. Rozenberg (Ed.), Handbook of Graph Grammars and Computing by Graph

Transformations, Vol. 1, World Scientific, Singapore, 1997, pp. 313–400. MR1480957
12. B. Courcelle, P. Weil, The recognizability of sets of graphs is a robust property, Theoret. Comput. Sci. to appear. cf. MR 2006d:68186
13. K. Denecke, S.L. Wismath, Universal Algebra and Applications in Theoretical Computer Science, Chapman & Hall, New York, 2002. MR1887177 (2003m:08001)
14. V. Diekert, Combinatorics on Traces, Lecture Notes in Computer Science, Vol. 454, Springer, Berlin, 1987. MR1075995 (92d:68070)
15. J. Doner, Tree acceptors and some of their applications, J. Comput. System Sci. 4 (1970) 406–451. MR0287977 (44 #5179a)
16. S. Eilenberg, Automata, Languages, and Machines, Vols. A and B, Academic Press, New York, 1974, 1976. MR0530382 (58 #26604a)
17. S. Eilenberg, J.B. Wright, Automata in general algebras, Inform. and Control 11 (1967) 452–470. MR0223285 (36 #6333)
18. C.C. Elgot, Decision problems of finite automata design and related arithmetics, Trans. Amer. Math. Soc. 98 (1961) 21–51. MR0139530 (25 #2962)
19. Z. Ésik, A variety theorem for trees and theories, Publ. Math. 54 (1999) 711–762. MR1709922 (2000h:68115)
20. Z. Ésik, P. Weil, On logically defined recognizable tree languages, in: P.K. Pandya, J. Radhakrishnan (Eds.), Proc. FST TCS 2003, Lecture Notes in Computer Science, Vol. 2914, Springer, Berlin, 2003, pp. 195–207. MR2093649 (2005e:68168)
21. Z. Ésik, P. Weil, Algebraic characterization of logically defined tree languages, in preparation.
22. D.M. Gabbay, A. Pnueli, S. Shelah, J. Stavi, On the temporal analysis of fairness, in: Proc. 12th ACM Symp. Principles of Programming Languages, Las Vegas, 1980, pp. 163–173.
23. F. Gécseg, M. Steiby, Tree Automata, Akadémiai Kiadó, Budapest, 1984. MR0735615 (86c:68061)
24. F. Gécseg, M. Steiby, Tree languages, in: G. Rozenberg, A. Salomaa (Eds.), Handbook of Formal Languages, Vol. 3, Springer, Berlin, 1997, pp. 1–68. MR1470018 (98m:68146)
25. G. Grätzer, Universal Algebra, Springer, Berlin, 1979. MR0538623 (80g:08001)
26. U. Heuter, First-order properties of trees, star-free expressions, and aperiodicity, in: R. Cori, M. Wirsing (Eds.), STACS 88, Lecture Notes in Computer Science, Vol. 294, Springer, Berlin, 1988, pp. 136–148. MR0935795 (89c:68058)
27. J.A. Kamp, Tense logic and the theory of linear order, Ph.D. Thesis, UCLA, 1968.
28. S. MacLane, Categories for the Working Mathematician, Springer, Berlin, 1971. MR0354798 (50 #7275)
29. R. McNaughton, S. Papert, Counter-Free Automata, MIT Press, Cambridge, MA, 1971. MR0371538 (51 #7756)
30. J. Mezei, J.B. Wright, Algebraic automata and context-free sets, Inform. and Control 11 (1967) 3–29. MR0234775 (38 #3091)
31. J.-E. Pin, Variétés de langages formels, Masson, Paris, 1984 (English translation: Varieties of formal languages, Plenum, New York, 1986). MR0752695 (86a:68058)
32. A. Potthoff, Modulo counting quantifiers over finite trees, in: J.-C. Raoult (Ed.), CAAP ’92, Lecture Notes in Computer Science, Vol. 581, Springer, Berlin, 1992, pp. 265–278.

MR1252005 (94g:03020)
33. A. Potthoff, First order logic on finite trees, in: P.D. Mosses, M. Nielsen, M.I. Schwartzbach (Eds.), TAPSOFT '95, Lecture Notes in Computer Science, Vol. 915, Springer, Berlin, 1995, pp. 125–139.
34. M.P. Schützenberger, On finite monoids having only trivial subgroups, Inform. and Control 8 (1965) 190–194. MR0176883 (31 #1154)
35. M. Steinby, A theory of tree language varieties, in: M. Nivat, A. Podelski (Eds.), Tree Automata and Languages, North-Holland, Amsterdam, 1992, pp. 57–81. MR1196732 (94a:68076)
36. M. Steinby, General varieties of tree languages, Theoret. Comput. Sci. 205 (1998) 1–43. MR1638624 (99m:68111)
37. H. Straubing, Finite Automata, Formal Logic, and Circuit Complexity, Birkhaüser, Boston, MA, 1994. MR1269544 (95f:68173)
38. J.W. Thatcher, J.B. Wright, Generalized finite automata theory with an application to a decision problem of second-order logic, Math. Systems Theory 2 (1968) 57–81. MR0224476 (37 #75)
39. W. Wechler, Universal algebra, EATCS Monographs on Theoretical Computer Science, Vol. 10, Springer, Berlin, 1992. MR1177406 (94a:08001)
40. P. Weil, Algebraic recognizability of languages, in: J. Fiala, V. Koubek, J. Kratochvíl (Eds.), MFCS 2004, Lecture Notes in Computer Science, Vol. 3153, Springer, Berlin, 2004, pp. 149–175. MR2143137 (2005m:68149)
41. Th. Wilke, An algebraic characterization of frontier testable tree languages, Theoret. Comput. Sci. 154 (1996) 85–106. MR1374381 (97f:68134)

Note: This list reflects references listed in the original paper as accurately as possible with no attempt to correct errors.