Zbl 1092.68060

Steinby, Magnus

Algebraic classifications of regular tree languages. (English)

Kudryavtsev, Valery B. et al., Structural theory of automata, semigroups, and universal algebra. Proceedings of the NATO Advanced Study Institute, Montreal, Quebec, Canada, July 7–18, 2003. Dordrecht: Kluwer Academic Publishers. NATO Science Series II: Mathematics, Physics and Chemistry 207, 381-432 (2005). ISBN 1-4020-3815-1/hbk; ISBN 1-4020-3816-X/pbk; ISBN 1-4020-3817-8/e-book

This is another excellent survey paper by the author, subsequent to his joint paper with F. Gécseg [*G. Rozenberg* and *A. Salomaa*, Tree lanuages, in: Handbook of Formal Languages, Vol. 3, (Berlin): Springer. not consec. pag. (1997; Zbl 0866.68057)] (1-69), with more up-to-date information and more emphasize on the algebraic theory of the tree languages.

After giving a thorough introduction in Section 1 and presenting algebraic preliminaries in Section 2, the author reviews finite automata and regular string languages in Section 3. Then trees are defined as terms in Section 4. Let Σ be a ranked alphabet, i.e., any $\sigma \in \Sigma$ has a non-negative integer rank which is interpreted as the arity of the functions that σ represents. For any finite alphabet X disjoint from Σ, called a leaf alphabet, ΣX-trees are defined as Σ-terms over X. Then universal algebra becomes a natural mathematical framework for the theory of tree automata and tree languages. Thus in Section 5, where tree recognizers are defined, we see that Σ-algebras can recognize ΣX-tree languages. Recognizable tree languages are sets of trees that can be recognized by finite tree automata (algebras). In Section 6 some operations on tree languages are introduced, and it is shown that the family of all recognizable tree languages is closed under many of them. A variety of tree languages is a family of tree languages closed under Boolean operations, inverse translations and inverse morphisms, as defined in Section 7. In Section 8 syntactic algebras and syntactic congruences of tree languages are introduced. The syntactic algebra of a tree language is the minimal algebra that recognizes the language (and it is finite iff the language is regular) and the syntactic congruence of a tree language is the greatest congruence relation on the term algebra that saturates the language (and it is of finite index iff the language is regular). In Section 9 varieties of finite algebras are defined to be classes of finite Σ-algebras closed under subalgebras, homomorphic images, and finite direct products. In Section 10 varieties of finite congruence relations are defined as families of filters of congruences of finite index on term algebras closed under inverse homomorphisms. These three kinds of varieties (of tree languages, algebras, and congruences) are linked together in Section 11 in the variety theorem which constitutes the central theme of the paper. It generalizes Eilenberg's (and Thérien's) variety theorem(s) from words to trees, and was proved by the author and independently by J. Almeida. It turns out that many of the known families of special regular tree languages are varieties of tree languages. In Section 12 several examples of varieties of tree languages and instances of the variety theorem are presented. In Section 13 two main generalizations of the variety theory are considered. Firstly, recognizable tree languages can be replaced with

recognizable subsets of free algebras over a give fixed variety of algebras **V**. Then we have the variety theorem for these families of recognizable sets and varieties of finite algebras contained in **V**. Secondly, we can consider families of tree languages and classes of algebras not necessary over a fixed ranked alphabet. This leads to the author's generalized variety theory introduced in 1998. In Section 14 some alternative approaches to classifications of tree languages are reviewed, such as the monoid/semigroup approach introduced by W. Thomas, the theories by Z. Ésik, the pre-clones of Z. Ésik and P. Weil, and tree algebras introduced by T. Wilke. Also the reviewer's variety theorem for monoids/semigroups and the variety theorem for tree algebras proved jointly by the reviewer and the author are mentioned. The last section (15) is devoted to the tree languages recognized by deterministic top-down (or root-to-frontier) tree recognizers, and the syntactic path monoids/semigroups introduced by F. Gécseg and the author are proposed as suitable syntactic invariants for characterizing families of DT-recognizable tree languages. The paper concludes with an extensive bibliography which can be useful for anybody entering this fascinating field.
The survey is very well-organized and contains most of the full proofs presented in a rigorous way, and it welcomes the newcomers who are well informed about string languages and want to explore the realm of trees. It can also be used by people working in term rewriting and (finite) universal algebra.

Saeed Salehí (Zanjan)

Keywords : Tree automata and tree languages; Varieties of finite algebras; Varieties of tree languages; Universal algebra
Classification :
∗68Q70 Algebraic theory of automata
08A70 Appl. of universal algebra in computer science