

Some Results of Zoltán Ésik on Regular Languages

Jean-Éric Pin^(✉)

IRIF, CNRS, Paris, France
Jean-Eric.Pin@irif.fr

Abstract. Zoltán Ésik published 2 books as an author, 32 books as editor and over 250 scientific papers in journals, chapters and conferences. It was of course impossible to survey such an impressive list of results and in this lecture, I will only focus on a very small portion of Zoltán's scientific work. The first topic will be a result from 1998, obtained by Zoltán jointly with Imre Simon, in which he solved a twenty year old conjecture on the shuffle operation. The second topic will be his algebraic study of various fragments of logic on words. Finally I will briefly describe some results on commutative languages obtained by Zoltán, Jorge Almeida and myself.

1 Regular Languages and Varieties

Let A be a finite alphabet. Let L be a language of A^* and let x and y be words of A^* . The *quotient* $x^{-1}Ly^{-1}$ of L by x and y is defined by the formula

$$x^{-1}Ly^{-1} = \{u \in A^* \mid xuy \in L\}.$$

A *lattice of languages* of A^* is a set \mathcal{L} of languages of A^* containing \emptyset and A^* and closed under finite union and finite intersection. It is *closed under quotients* if every quotient of a member of \mathcal{L} is also in \mathcal{L} . A *Boolean algebra of languages* is a lattice of languages closed under complement.

A *class of languages* is a correspondence \mathcal{C} which associates with each alphabet A a set $\mathcal{C}(A^*)$ of regular languages of A^* . A *variety of languages* is a class of languages \mathcal{V} closed under Boolean operations, quotients and inverses of morphisms. This means that, for each alphabet A , $\mathcal{V}(A^*)$ is a Boolean algebra of regular languages closed under quotients and that if $\varphi : A^* \rightarrow B^*$ is a monoid morphism, then $L \in \mathcal{V}(B^*)$ implies $\varphi^{-1}(L) \in \mathcal{V}(A^*)$.

A *variety of finite monoids* is a class of finite monoids which is closed under taking submonoids, homomorphic images and finite direct products. Eilenberg's variety theorem [6] gives a bijective correspondence between varieties of monoids and varieties of languages. Let \mathbf{V} be a variety of finite monoids and, for each alphabet A , let $\mathcal{V}(A^*)$ be the set of all languages of A^* whose syntactic monoid is in \mathbf{V} . Then \mathcal{V} is a variety of languages. Furthermore, the correspondence $\mathbf{V} \rightarrow \mathcal{V}$ is a bijection between varieties of monoids and varieties of languages.

A language L is *commutative* if any word obtained by permuting the letters of a word of L also belongs to L . A variety of languages is *commutative* if all of its languages are commutative, or equivalently, if all the monoids of the corresponding variety of finite monoids are commutative.

A *renaming* or *length-preserving morphism* is a morphism φ from A^* into B^* , such that, for each word u , the words u and $\varphi(u)$ have the same length. It is equivalent to require that, for each letter a , $\varphi(a)$ is also a letter, that is, $\varphi(A) \subseteq B$. Similarly, a morphism is *length-decreasing* if the image of each letter is either a letter or the empty word. Finally, a morphism is *length-multiplying* if all letters have images of the same length, that is, if $\varphi(A) \subseteq B^k$ for some integer k .

2 The Shuffle Operation

Recall that the *shuffle product* (or simply *shuffle*) of two languages L_1 and L_2 over A is the language

$$L_1 \sqcup L_2 = \{w \in A^* \mid w = u_1v_1 \cdots u_nv_n \text{ for some words } u_1, \dots, u_n, \\ v_1, \dots, v_n \text{ of } A^* \text{ such that } u_1 \cdots u_n \in L_1 \text{ and } v_1 \cdots v_n \in L_2\}.$$

The shuffle product defines a commutative and associative operation over the set of languages over A .

Zoltán Ésik has long been interested in the shuffle operation. In a series of papers with Bertol or with Bloom [3–5, 8, 9], he studied the free shuffle algebra and proved that the equational theory of shuffle has no finite axiomatization. This was probably the reason why he got interested in the conjecture proposed by Perrot [12] in 1978.

Perrot wanted to characterize the varieties of languages closed under shuffle. He was able to characterize all commutative varieties closed under shuffle, but failed to characterize the noncommutative ones. However, he conjectured that the only noncommutative variety of languages closed under shuffle is the variety of all regular languages.

This conjecture remained open for twenty years, until Zoltán Ésik, in collaboration with another famous Hungarian-born computer scientist, Imre Simon, managed to solve the conjecture positively [11]. Their proof is very ingenious and was the starting point of further development.

3 Logic on Words

In December 2001, Ésik and Ito published a BRICS report entitled *Temporal logic with cyclic counting and the degree of aperiodicity of finite automata*. In this paper, which was published in 2003 [10], the authors studied an enrichment of temporal logic involving cyclic counting and they provided an algebraic characterization of the corresponding class of regular languages. An instance of Eilenberg's variety theorem? Not quite, because this class is closed under inverses

of length-preserving morphisms, but is not closed under inverses of arbitrary morphisms.

The next year (2002), Ésik published another BRICS report, entitled *Extended temporal logic on finite words and wreath product of monoids with distinguished generators*, which became a DLT paper in 2003 [7]: in this paper, he further developed his idea of enriching temporal logic, in the spirit of Wolper [15]. He managed to give an algebraic characterization of several fragments by using monoids with distinguished generators. This led to a series of new results as well as a unified and elegant proof of known results.

It turns out that a similar idea was developed independently and at the same time by Straubing [14]. This gave rise to the theory of \mathcal{C} -varieties, which is an extension of Eilenberg's variety theory. The letter \mathcal{C} refers to a class of morphisms (called \mathcal{C} -morphisms) between free monoids (for instance *length-preserving* morphisms, *length-decreasing* or *length-multiplying* morphisms). Now a \mathcal{C} -variety of languages is defined as a *variety of languages* except that it is only closed under inverses of \mathcal{C} -morphisms.

The corresponding algebraic objects are no longer monoids, but *stamps*, that are surjective monoid morphisms $\varphi : A^* \rightarrow M$ from a finitely generated free monoid A^* onto a finite monoid M .

4 Back to the Shuffle Operation

In 1995, I proposed another extension of Eilenberg's variety theorem [13]. A positive variety of languages is defined exactly like a variety of languages, except that it is not closed under complement. In other words, for each alphabet A , $\mathcal{V}(A^*)$ is not required to be a Boolean algebra of languages, but only a lattice of languages. For the algebraic counterpart, one needs to consider ordered monoids instead of monoids.

The question now arises to describe all positive varieties closed under shuffle. Some progress in this direction can be found in [1], but the first step, namely the commutative case, was clarified only recently, in a paper of Jorge Almeida, Zoltán Ésik and myself [2].

I am sure that Zoltán would have liked to further investigate this type of questions among the numerous topics he was interested in. I deeply miss him, as a scientist and as a personal friend.

References

1. Almeida, J., Cano, A., Klíma, O., Pin, J.-É.: Fixed points of the lower set operator. *Int. J. Algebra Comput.* **25**(1–2), 259–292 (2015)
2. Almeida, J., Ésik, Z., Pin, J.-É.: Commutative positive varieties of languages. *Acta Cybern.* **23**, 91–111 (2017)
3. Bloom, S.L., Ésik, Z.: Free shuffle algebras in language varieties extended abstract. In: Baeza-Yates, R., Goles, E., Pobleto, P.V. (eds.) *LATIN 1995*. LNCS, vol. 911, pp. 99–111. Springer, Heidelberg (1995). doi:[10.1007/3-540-59175-3_84](https://doi.org/10.1007/3-540-59175-3_84)

4. Bloom, S.L., Ésik, Z.: Nonfinite axiomatizability of shuffle inequalities. In: Mosses, P.D., Nielsen, M., Schwartzbach, M.I. (eds.) CAAP 1995. LNCS, vol. 915, pp. 318–333. Springer, Heidelberg (1995). doi:[10.1007/3-540-59293-8_204](https://doi.org/10.1007/3-540-59293-8_204)
5. Bloom, S.L., Ésik, Z.: Free shuffle algebras in language varieties. Theoret. Comput. Sci. **163**(1–2), 55–98 (1996)
6. Eilenberg, S.: Automata, Languages and Machines, vol. B. Academic Press, New York (1976)
7. Ésik, Z.: Extended temporal logic on finite words and wreath product of monoids with distinguished generators. In: Ito, M., Toyama, M. (eds.) DLT 2002. LNCS, vol. 2450, pp. 43–58. Springer, Heidelberg (2003). doi:[10.1007/3-540-45005-X_4](https://doi.org/10.1007/3-540-45005-X_4)
8. Ésik, Z., Bertol, M.: Nonfinite axiomatizability of the equational theory of shuffle. In: Fülöp, Z., Gécseg, F. (eds.) ICALP 1995. LNCS, vol. 944, pp. 27–38. Springer, Heidelberg (1995). doi:[10.1007/3-540-60084-1_60](https://doi.org/10.1007/3-540-60084-1_60)
9. Ésik, Z., Bertol, M.: Nonfinite axiomatizability of the equational theory of shuffle. Acta Inform. **35**(6), 505–539 (1998)
10. Ésik, Z., Ito, M.: Temporal logic with cyclic counting and the degree of aperiodicity of finite automata. Acta Cybern. **16**, 1–28 (2003)
11. Ésik, Z., Simon, I.: Modeling literal morphisms by shuffle. Semigroup Forum **56**, 225–227 (1998)
12. Perrot, J.-F.: Variétés de langages et opérations. Theoret. Comput. Sci. **7**, 197–210 (1978)
13. Pin, J.-E.: A variety theorem without complementation. Russ. Math. (Iz. VUZ) **39**, 80–90 (1995)
14. Straubing, H.: On logical descriptions of regular languages. In: Rajsbaum, S. (ed.) LATIN 2002. LNCS, vol. 2286, pp. 528–538. Springer, Heidelberg (2002). doi:[10.1007/3-540-45995-2_46](https://doi.org/10.1007/3-540-45995-2_46)
15. Wolper, P.: Temporal logic can be more expressive. Inform. Control **56**(1–2), 72–99 (1983)