

A bibliography on Szilard languages

Erkki Mäkinen

*Department of Computer and Information Sciences, P.O. Box 607,
FIN-33014 University of Tampere, Finland*

Abstract

This note lists the Szilard language literature so far appeared. Both Chomsky type and various other types of grammars are considered.

Keywords: formal languages, Szilard languages.

1 Introduction

The derivational structure of a grammar can be studied by considering the language obtained by first uniquely labeling the productions and then recording the strings of labeled productions used in terminal derivations (when applicable) of the grammar. The language so obtained is usually called 'Szilard language' although the names 'label language', 'derivation language' and 'associate language' are also used in the literature. Throughout this paper we use the phrase 'Szilard language'.

We first define the Szilard language of a normal (Chomsky type) context-free grammar $G = (V, \Sigma, P, S)$ as follows. Suppose the productions are uniquely labeled with the symbols of an alphabet C . If a production $A \rightarrow \alpha$ is associated with a label ρ , we write $\rho : A \rightarrow \alpha$. If a sequence $\rho = \rho_1\rho_2 \dots \rho_n$ of labeled productions is applied in a derivation $\beta \Rightarrow^* \gamma$, we can write $\beta \Rightarrow^\rho \gamma$. The *Szilard language* $Sz(G)$ of G is defined as

$$Sz(G) = \{\phi \mid S \Rightarrow^\phi w, w \in \Sigma^*\}.$$

Hence, $Sz(G)$ is a language over the alphabet C used in labeling productions.

The *left Szilard language* of a context-free grammar is obtained if the relation \Rightarrow above is replaced by the relation \Rightarrow_l . The basic idea of Szilard languages as a tool for describing the derivational structure of a grammar can be applied in various kinds of grammar types (see chapter 4).

2 Context-free grammars

Early contributions in the field are due to Altman and Banerji [1] and Stotskij [61], but systematic work on Szilard languages started at the early 1970's [10,11,16,54,57,60].

For example, the following aspects of Szilard languages and left Szilard languages of context-free grammars are studied:

- regularity and context-freeness [29,34,54,57]
- recognition and tape complexity [19,20,58]
- various decidability problems [34,39,49,57,62]
- boundedness [27,28]
- the form of homomorphic images [18,36,43,56].

On the other hand, Szilard languages can be considered as examples of some interesting language classes, such as permutative languages [35]. Similarly, most left Szilard languages of context-free grammars are pure [38].

Szilard languages can naturally be studied also in the connection with other context-free derivations than general or leftmost derivations [40,42] and in the connection with more general Chomsky type grammars than context-free grammars [13,14,54].

A somewhat similar concept to Szilard languages is that of control sets [12]. Control sets are not considered in this bibliography.

3 Applications

Szilard languages have some applications outside the borders of formal language theory.

Compressing a program file can be efficiently done by noticing the derivational structure of the programming language used. The structure of the programming language in question is natural to express by using the left Szilard language [23].

Grammar transformations and translations are an obvious area of applying (left) Szilard languages. These questions are studied by Kriegel and Maurer [25], Kriegel and Ottmann [26], and Mäkinen [30,31].

Since left Szilard languages are deterministic in a certain very strict sense, they can be shown to be inferable in the limit from positive data only [64].

Simpler algorithms are known for the Szilard languages of regular grammars [41,46,47]. The ability to infer the left Szilard language from positive data only can also be used when characterizing the complexity of a context-free grammar [45]. Szilard languages are also applied in grammatical inference by Firoiu et al. [9].

4 Other grammar types

Szilard languages are defined at least for the following grammar types:

- label grammars [22]
- pure grammars [37,38]
- matrix grammars [15,55]
- L-systems [52,63]
- IO-grammars [8]
- parallel communicating grammar systems [53]
- co-operating distributed grammar systems [6].

The relationship between Szilard languages and Petri net languages is also studied in the literature [5,21].

5 Conclusions

The bibliography given below is most likely incomplete. The author welcomes supplementing information by e-mail (em@cs.uta.fi). Updated versions of this bibliography will be available in <ftp://cs.uta.fi/pub/reports/Szilard.ps.Z> and <ftp://cs.uta.fi/pub/reports/pdf/Szilard.pdf>.

Notice that all the references listed below are not cited in the text.

References

- [1] E. Altman and R. Banerji, Some problems of finite representability. *Inform. Control* **8** (1965), 251–263.
- [2] A.T. Borgida, Some formal results about stratificational grammars and their relevance to linguistics. *Math. Systems Theory* **16** (1983), 29–56.
- [3] A. Burago, Learning structurally reversible context-free grammars from queries and counterexamples in polynomial time. In. *Proc. COLT '94*, ACM, 1994, 140–146.

- [4] P. Butzbach, Une famille de congruences de Thue pour lesquelles le problème de l'équivalence est décidable. Application à l'équivalence des grammaires séparées. In: M. Nivat (ed.), *Automata, Languages and Programming*, North-Holland, 1973, 3–12.
- [5] S. Crespi-Reghizzi and D. Mandrioli, Petri nets and Szilard languages. *Inform. Control* **33** (1977), 177–192.
- [6] J. Dassow, V. Mitrană and Gh. Păun, Szilard languages associated to co-operating distributed grammar systems. *Stud. Cercet. Mat.* **45** (1993), 403–413.
- [7] J. Dassow and Gh. Păun, *Regulated Rewriting in Formal Language Theory*. Springer, 1989.
- [8] J. Duske, R. Parchmann and J. Specht, Szilard languages of IO-grammars. *Inform. Control* **40,3** (1979), 319–331.
- [9] L. Firoiu, T. Oates P.R. and Cohen, Paul R. Learning regular languages from positive evidence. In Proc. of the Twentieth Annual Meeting of the Cognitive Science Society, 1998, pp. 350–355.
- [10] A.C. Fleck, On the combinatorial complexity of context-free grammars. *Proceedings of IFIP Conference TA 2* (1971), 15–17.
- [11] A.C. Fleck, An analysis of grammars by their derivation sets. *Inform. Control* **24** (1974), 389–398.
- [12] S. Ginsburg and E.H. Spanier, Control sets on grammars. *Math. Systems Theory* **2** (1968), 159–177.
- [13] J. M. Hart, Derivation languages and syntactical categories. *Inform. Control* **28**, (1975), 204–220.
- [14] J. M. Hart, The derivation language of a phrase structure grammar. *J. Comput. System Sci.* **12**, 1 (1976), 64–79.
- [15] M. Höpner, Über den Zusammenhang von Szilardsprachen und Matrixgrammatiken. Bericht Nr. 12, Institut für Informatik, Universität Hamburg, 1974.
- [16] M. Höpner, Eine Charakterisierung der Szilardsprachen und ihre Verwendung als Steuersprachen. *Lecture Notes in Computer Science* **26** (1974), 113–121.
- [17] M. Höpner and M. Opp, About three classes of languages built up by shuffle operations. *Lecture Notes in Computer Science* **45** (1976), 337–344.
- [18] M. Höpner and M. Opp, Renaming and erasing in Szilard languages. *Lecture Notes in Computer Science* **52** (1977), 244–257.
- [19] O.H. Ibarra, S.K. Kim and L.E. Rosier, Some characterizations of multihead finite automata. *Inform. Control* **67**, (1985), 114–125.
- [20] Y. Igarashi, The tape complexity of some classes of Szilard languages. *SIAM J. Comput.* **6**, 3 (1977), 460–466.

- [21] M. Jantzen, On the hierarchy of Petri net languages. *R.A.I.R.O. Informatique théorique* **13**, 1 (1979), 19–30.
- [22] M. Jantzen and M. Opp, A normal form theorem for label grammars. *Math. Systems Theory* **14** (1981), 289–303.
- [23] J. Katajainen, M. Penttonen and J. Teuhola, Syntax-directed compression of program files. *Software-Practice and Experience* **16**, 3 (1983), 269–276.
- [24] T. Koshiba, On a hierarchy of slender languages based on control sets. *Fundam. Inf.* **31**, 1 (1997), 41–47.
- [25] H.P. Kriegel and H.A. Maurer, Formal translations and Szilard languages. *Inform. Control* **30** (1976), 187–198.
- [26] H.P. Kriegel and Th. Ottmann, Left-fitting translations. *Lecture Notes in Computer Science* **52** (1977), 309–322.
- [27] E. Mäkinen, On certain properties of left Szilard languages. *Elektron. Inf. Verarb. Kybern. EIK* **19**, 10/11 (1983), 497–501.
- [28] E. Mäkinen, Boundedness testing for unambiguous context-free grammars. *Inform. Process. Lett.* **17**, (1983), 181–183.
- [29] E. Mäkinen, On context-free and Szilard languages. *BIT* **24**, (1984), 164–170.
- [30] E. Mäkinen, On derivation preservation. *Inform. Process. Lett.* **19**, (1984), 225–228.
- [31] E. Mäkinen, A note on undercover relation. *Inform. Process. Lett.* **20**, (1985), 19–21.
- [32] E. Mäkinen, A note on depth-first derivations. *BIT* **25**, (1985), 293–296.
- [33] E. Mäkinen, An undecidable problem for context-free grammars. *Inform. Process. Lett.* **20**, (1985), 141–142.
- [34] E. Mäkinen, On context-free derivations. *Acta Universitatis Tamperensis, Ser. A.* **198**, Tampere 1985.
- [35] E. Mäkinen, On permutative grammars generating context-free languages. *BIT* **25**, (1985), 604–610.
- [36] E. Mäkinen, On homomorphic images of Szilard languages. *Intern. J. Computer Mathematics* **18**, (1986), 239–245.
- [37] E. Mäkinen, On Szilard languages of pure context-free grammars. *Elektron. Inf. Verarb. Kybern. EIK* **22**, 10/11 (1986), 527–532.
- [38] E. Mäkinen, A note on pure grammars. *Inform. Process. Lett.* **23**, (1986), 271–274.
- [39] E. Mäkinen, A note on the inclusion problem for Szilard languages. *Intern. J. Computer Mathematics* **21**, (1987), 291–295.

- [40] E. Mäkinen, On breadth-first context-free grammars. *J. Inf. Process. EIK* **26**, 3 (1990), 129–135.
- [41] E. Mäkinen, The grammatical inference problem for the Szilard languages of linear grammars. *Inform. Process. Lett.* **36**, (1990), 203–206.
- [42] E. Mäkinen, A hierarchy of context-free derivations. *Fundam. Inf.* **14**, 2 (1991), 255–259.
- [43] E. Mäkinen, On homomorphic images of left Szilard languages. *Intern. J. Computer Mathematics* **46**, (1992), 131–135.
- [44] E. Mäkinen, On grammatical inference and derivational complexity of context-free grammars. *Fundam. Inf.* **17**, 4 (1992), 363–368.
- [45] E. Mäkinen, Remarks on the structural grammatical inference for context-free grammars. *Inform. Process. Lett.* **44**, (1992), 125–127.
- [46] E. Mäkinen, A note on the grammatical inference problem for even linear languages. *Fundam. Inf.* **25**, 2 (1996), 175–181.
- [47] E. Mäkinen, A family of languages which is polynomial-time learnable from positive data in Pitt’s sense. *Intern. J. Computer Mathematics* **61**, (1996), 175–179.
- [48] E. Mäkinen, Ranking and unranking left Szilard languages. *Intern. J. Computer Mathematics* **68**, (1998), 29–38.
- [49] E. Mäkinen, On the inclusion problem for very simple deterministic pushdown automata. *Journal of Automata, Languages and Combinatorics* **6**, (2001), 351–357.
- [50] A. Mateescu and A. Salomaa, Aspects of classical language theory. In: G. Rozenberg and A. Salomaa (eds.), *Handbook of Formal Languages. Vol. 1*, Springer, 1997, 175–251.
- [51] H.A. Maurer, A. Salomaa and D. Wood, Derivation languages of grammar forms. *Intern. J. Computer Mathematics* **9**, (1981), 117–130.
- [52] H.A. Maurer and D. Wood, Szilard languages of L-systems. In: A.Lindenmayer and G. Rozenberg (eds.) *Automata, Languages, Development*, North-Holland, 1976, 251–261.
- [53] V. Mihalache, Szilard languages associated to parallel communicating grammar systems. Turku Centre for Computer Science, Finland, Tech. Report **TUCS-TR-14**, May 1996.
- [54] E. Moriya, Associate languages and derivational complexity of formal grammars and langauges. *Inform. Control* **22** (1973), 139–162.
- [55] Gh. Păun, On Szilard’s languages associated to a matrix grammar. *Inform. Process. Lett.* **8**, 2 (1979), 104–105.

- [56] Gh. Păun, On some families of Szilard languages. *Bull. Math. de la Soc. Sci. Math. De la R.S. de Roumanie* **27(75)**, 3 (1983), 259–265.
- [57] M. Penttonen, On derivation languages corresponding to contex-free grammars. *Acta Inform.* **3** (1974), 285–291.
- [58] M. Penttonen, Szilard languages are log n tape recognizable. *Elektron. Inf. Verarb. Kybern. EIK* **13**, 11 (1977), 595–602.
- [59] G.E. Révész, *Introduction to Formal Languages*. McGraw-Hill, 1980.
- [60] A. Salomaa, *Formal Languages*. Academic Press, 1973.
- [61] E.D. Stotskij, Some restrictions on derivations in context-free grammars. *Nauchno-Techn. Inform. Ser.* **2**, 7 (1967), 35–38.
- [62] M. Wakatsuki and E. Tomita, A fast algorithm for checking the inclusion for very simple deterministic pushdown automata. *IEICE Trans. on Information and Systems* **E76-D**, 10 (1993), 1224–1233.
- [63] D. Wood, A note on Lindenmayer systems, Szilard languages spectra, and equivalence. *International Journal of Computer and Information Sciences* **4**, 1 (1975), 53–62.
- [64] T. Yokomori, Polynomial time learning of very simple grammars from positive data. *Proc. 4th Workshop on Computational Learning Theory* (1991), 213–227.