

Erkki Mäkinen

**State change languages as
homomorphic images of
Szilard languages**



UNIVERSITY OF TAMPERE
SCHOOL OF INFORMATION SCIENCES
REPORTS IN INFORMATION SCIENCES 48

TAMPERE 2016

UNIVERSITY OF TAMPERE
SCHOOL OF INFORMATION SCIENCES
REPORTS IN INFORMATION SCIENCES 48
NOVEMBER 2016

Erkki Mäkinen

**State change languages as
homomorphic images of
Szilard languages**

SCHOOL OF INFORMATION SCIENCES
FIN-33014 UNIVERSITY OF TAMPERE

ISBN 978-952-03-0322-8

ISSN-L 1799-8158
ISSN 1799-8158

STATE CHANGE LANGUAGES AS HOMOMORPHIC IMAGES OF SZILARD LANGUAGES

ERKKI MÄKINEN

Çevik and Kiliç have introduced a class of regular languages which imitate the function of finite automata accepting regular languages: When an automaton changes its state the transition indicating the change is labeled with 1 and when the automaton keeps its state the self-loop transition is labeled with 0. Regular languages over $\{0, 1\}^*$ accepted by such automata are called state change languages.

We consider state change languages in the context of Szilard languages of regular grammars. We show that state change languages are homomorphic images of Szilard languages and that other similar classes of state change type of languages can be defined.

Keywords: regular grammar, Szilard language

Classification: 68Q45

1. INTRODUCTION

Çevik and Kiliç [1] have introduced a class of regular languages which imitate the function of finite automata accepting regular languages: When an automaton changes its state the transition indicating the change is labeled with 1 and when the automaton keeps its state the self-loop transition is labeled with 0. Regular languages over $\{0, 1\}^*$ accepted by such automata are called state change languages.

In this note we consider state change languages in the context of Szilard languages of regular grammars and establish a hierarchy of grammars which model the derivational structure of regular grammars in different levels of accuracy. State change languages describe the derivational structure in a coarse level showing only when changes of states have happened but even the sources or goals of the changes are ignored. More accurate picture of the derivational structure is obtained when we record the states in which the control of the computation is in different times during the computation. This is done with so called nonterminal language. Finally, a complete picture of the derivational structure of a regular grammar is obtained by considering its Szilard language.

2. PRELIMINARIES

If not otherwise stated we follow the notations and definitions of [2]. Let $G = (N, T, P, S)$ be a regular grammar where N is the alphabet of nonterminals, T is the alphabet of terminals, P is the set of productions, and S is the start symbol.

Suppose the productions of G 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 *Szilar 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 the productions [3].

Clearly, the Szilar language $Sz(G)$ of a regular grammar G is itself a regular language. The finite automaton accepting $Sz(G)$ has the property that each transition is uniquely labeled. The unique labeling of transitions corresponds the unique labeling of productions in G .

3. STATE CHANGE LANGUAGES

Çevik and Kiliç [1] started with a finite automaton A accepting a regular language R . The labels of the transition in A are changed so that each transitions traversing from a state to another state is labeled with 1 and each transition traversing from a state to the same state (a self-loop) is labeled with 0. The language accepted by the new automaton is the state change language $SC(A)$ of A . The properties of the class of languages accepted by such transformed automata are studied by Çevik and Kiliç [1].

Let $G = (N, T, P, S)$ be a regular grammar whose productions are uniquely labeled with the symbols of an alphabet C . Consider a homomorphism f from C to $\{0, 1\}$ defined such that if $\rho : A \rightarrow aA$, then $f(\rho) = 0$, and otherwise, i.e., ρ is the label of a production of type $\rho : A \rightarrow aB$, $A \neq B$, $a \in T \cup \{\lambda\}$, or of type $\rho : A \rightarrow a$, $a \in T \cup \{\lambda\}$, then $f(\rho) = 1$. If A is the finite automaton obtained from G by the standard manner in which a state in the automaton corresponds to a nonterminal in the grammar, then $f(Sz(G)) = SC(A)$.

Similarly, we can define a homomorphism h from C to T by letting $f(\rho) = a$ if a is the terminal symbol in the right hand side of the production labeled with ρ . If productions with no terminal symbols in the right hand side are allowed, then $h(\rho) = \lambda$, when the production labeled with ρ has the form $A \rightarrow B$ or $A \rightarrow \lambda$. Now we have $h(Sz(G)) = L(G)$.

Still another natural class of regular languages can be defined in the same manner. Namely, we can take into account the state from which the control of the automaton has left, or equivalently, the nonterminal to which the production has been applied. This gives us a homomorphism g from C to N defined by $g(\rho) = A$, if ρ is the label of a production with nonterminal A in its left-hand side. The language $g(Sz(G)) = NL(G)$ is called the nonterminal language of G . (Notice that a similar homomorphism is defined in [4] for the context-free grammars and called leftmost nonterminal language.)

The automaton construction related to $NL(G)$ has the property that all transitions leaving a state have the same label. Similarly, we could consider the case where all the

transitions entering a state have the same label. However, we omit this construction here.

Hence, we have the situation depicted in Figure 1.

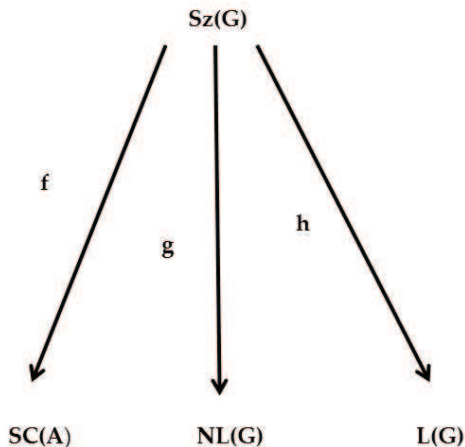


Fig. 1. Three homomorphic images of $Sz(G)$.

4. SOME REMARKS ON NONTERMINAL LANGUAGES AND THEIR RELATIONSHIP BETWEEN STATE CHANGE LANGUAGES

Consider a regular grammar G and its nonterminal language $NL(G)$. As each derivation in G starts from the unique start symbol, say S , each word in $NL(G)$ start with S .

Let G be a regular grammar, A the corresponding finite automaton, and $SC(A)$ its state change language. Each word in $SC(A)$ ends with 1, as the state must change in order to end the computation in A . In the grammar notation this means that the only nonterminal appearing in the sentential forms disappears.

If $NL(G)$ is the nonterminal language of G , we can define a mapping t from $NL(G)$ to $SC(A)$ as follows. If $a = a_1a_2 \dots a_n$ is a word in $NL(G)$ then $t(a) = b_1b_2 \dots b_{n-1}0$, where $b_i = 0$, if $a_i = a_{i+1}$ and $b_i = 1$, if $a_i \neq a_{i+1}$, for $i = 1, \dots, n - 1$.

Hence, we can augment Figure 1 with homomorphisms t as shown in Figure 2.

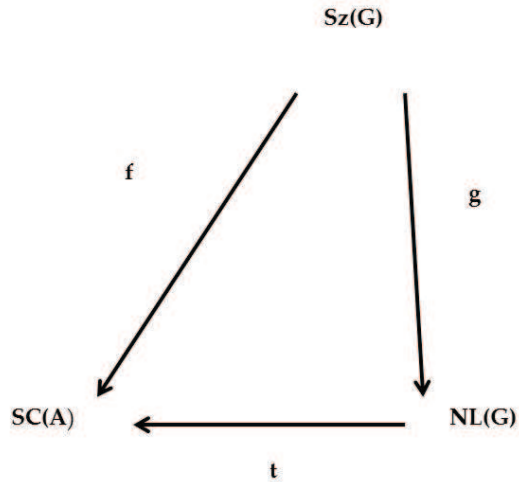


Fig. 2. The relationship between $SC(A)$ and $NL(G)$.

REFERENCES

- [1] A. Çevik and H. Kiliç: newblockAn infinite proper subset of regular languages as a atate change based coupling of finite automata. Proc. of the World Congress on Engineering and Computer Science 2015 *vol. 1*.
- [2] M.A. Harrison: Introduction to Formal Language Theory. Addison-Wesley, 1978.
- [3] E. Mäkinen: On context-free derivations. Acta Universitatis Tamperensis *Ser. A, Vol. 198, 1985*.
- [4] E. Mäkinen: On homomorphic images of left Szilard languages. Intern. J. Comput. Math. *46* (1992), 131–135.