International Journal of Engineering, Science and Mathematics
Vol. 8 Issue 5, May 2019,
ISSN: 2320-0294 Impact Factor: 6.765
Journal Homepage: http://www.ijesm.co.in, Email: ijesmj@gmail.com Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gage as well as in Cabell's Directories of Publishing Opportunities, U.S.A

SEMIGROUPS AND FORMAL LANGUAGES

Dr. Anand Jha

Deptt. of Mathematics

+2 Sarvoday High School, Darbhanga-846004

Email: anandajha75@gmail.com

Abstract

The underlying principle is always the same: Given an “alphabet” or “vocabulary” (consisting of letters, words, punctuation symbols, number symbols, programming instructions, etc.), we have a method (grammar) for constructing “meaningful” words or sentences from this alphabet. This immediately reminds us of the term “word semigroup” and indeed, these free semigroups will play a major role, the language L constructed will be a subset of the free semigroup A_* or the monoid A^* on the alphabet A.

Key words: Semigroup, Grammar, Automata.

Introduction:

Besides the natural languages such as English, so-called formal languages are of importance in the formal sciences, especially in computing or information science.

There are essentially three ways to construct a language L.

(i) Approach via grammar: given a collection of rules (grammar), generate L from A.

(ii) Approach via automata : consider an initial semi automation which processes the elements of L in a suitable way.

(iii) Algebraic approach : L is constructed by the algebraic combination of certain subsets of A.

All three approaches use algebraic method. Let us start with method (i) This approach is largely based on the work of Chomosky (1957).

1.1 DEFINITION:

A phase-structure grammar is a quadruple G = (A, G, $\rightarrow$, g_o), where A and G are none empty disjoint finite sets, $g_0 \in$ G and $\rightarrow$ is a finite relation from G into (A U G)*.

1.2 DEFINITION:

Let G be a grammar.

(i) A is called the alphabet of G.

(ii) G is called the set of grammar symbols.

(iii) V:= A U G is the complete vocabulary.

(iv) The elements (x, y) in $\rightarrow$ (which is a subset of G x V*) are also written in the form x $\rightarrow$ y and are called productions or rewriting rules.

(v) g_o is called the initial symbol.

We now obtain a suitable subset of A^*, by using G. In order to do this, we need another relation, this time on V^*.

1.3 DEFINITION:

Let $G = (A, G, \rightarrow, g_o)$ be a grammar. For $y, z \in V^*$, we write $y \Rightarrow z$ if there are $v \in G$ and $v_1, v_2, w \in V^*$ with

$Y = v_1 \, v \, v_2$, $z = v_1 \, w \, v_2$, and $v \rightarrow w$.

The reason for introducing $\Rightarrow$ is to obtain a new word $v_1, v_2 \ldots.. v'_r \ldots v_n$ from a given word $v_1 \, v_2 \ldots.. v_r \ldots\ldots v_n$ and a rules $v_r \rightarrow v'_r$. Thus we extend $\rightarrow$ to a compatible relation $\Rightarrow$ on V^*. Then transitive hull is $y \Rightarrow t$ of $\Rightarrow$ is given by $y \Rightarrow t \; z \Leftrightarrow$ there are $n \in N_o \; x_o, \ldots\ldots X_n \in V^*$

With $y = x_o \Rightarrow \ldots.. \Rightarrow X_n = z$.

The sequence $x_o, \ldots. X_n$ is called a derivation of z from y; n the length of the derivation.

1.4 DEFINITION:

Let $G = (A, G, \rightarrow, g_o)$ be a grammar

$L(G) := \{ l \in A^* | \; g_o \Rightarrow t \; l\}$

is called the language generated by G (also the phase-structure language).

1.5 DEFINITION:

A grammar $G =(A,G, \rightarrow,g_o)$ and is generated language are called right linear if all elements in $\rightarrow$ are of the forms $g \rightarrow a$ or $g \rightarrow ag'$. With $g, g' \in G$, $a \in A^*$. G is called context-free if for all $x \rightarrow y$ in $\rightarrow$ we have $l(x) = l$.

In a right linear language, the length of a word is never shortened by any derivation. All right linear grammars are context-free.

1.6 DEFINITION:

Let A be a set and $R \subseteq A^*$. R is called regular or a regular language if R can be obtained from one-element subset of A^* by a finite number of admissible operations. Admissible operations are the formations of unions, products, and generated submonoids.

The beginning of the study of formal languages can be traced to Chomsky, who in 1957 introduced the concept of a context-free language in order to model natural languages. Since the late 1960s there has been considerable activity in the theoretical development of context-free languages both in connection with natural languages and with the programming languages. Chomsky used Semi-Thue systems to define languages, which can be described as certain subjects of finitely generated free monoids. Chomsky (1957) details a revised approach in the light of experimental evidence and careful consideration of semantic and syntactic structure of sentences.

BBLIOGRAPHY :

1. Eilenberg, S (1974). Automata, Languages and Machines, Vols. I, II. New York : Academic Press.
2. Howie, J.M. (1995). Fundamentals of semigroup theory. Oxford : Clarendon Press.
3. Lidl, R. and G. Pilz (1984). Applied Abstract Algebra (1^{st} ed.). New York: Springer-Verlag. Russian edition: Ekaterinenburg, 1997.
4. Lidl, R. and H. Niederreiter (1994). Introduction to Finite Fields their Applications. Cambridge University Press.