FORMAL LANGUAGES

with Respect to the Chomsky-Schützenberger Hierarchy

Eric Kunze

27 January 2020

OUTLINE

1. Foundations of Formal Languages
2. Formal Grammar and the Chomsky-Hierarchy
3. Consequences and applications of Formal Languages

KEYWORDS

Formal Language, Formal Grammar, Chomsky-Schützenberger Hierarchy, Logic, Computability Theory

ABSTRACT

Every natural language complies with long time developed rules called grammar. A *formal language* describes the syntax of a language in a more formal way as a set of *words* over an (arbitrary) *alphabet*. In this way the structured construction, analysis and classification of languages are possible. In general, there are two methods of describing the syntax of words: generating strings by replacement starting with a fixed start symbol, or accepting given strings via an automaton. As Noam Chomsky in the 1950's did we focus on the former and analyse (formal) *grammar* connected with their generated languages. Afterwards we can classify languages by differentiating the corresponding grammar and especially their replacement behaviour. As result we get the so-called *Chomsky-Schützenberger Hierarchy* which distributes languages into four types: recursively enumerable, context-sensitive, context-free and regular languages, which are more restrictive in each step. This knowledge can be applied both in mathematics and computer science. Defining a language of logic, we get the fundament of (pure) mathematics and we can formalize terms like e.g. "proof" or "theorem". Furthermore, problems of computability theory can be specified and proven with the result that we are able to express the limits of computation.

VOCABULARY

Alphabet Σ: set of symbols

Language L: subset of words over an alphabet Σ ($L \subseteq \Sigma^*$)

Grammar $G = (N, \Sigma, P, S)$: Language description with an (finite) alphabet Σ (terminal symbols), a (finite) set N of non-terminal symbols, a set P of productions (replacement rules) and a start symbol $S \in N$

Production $A \to aAb$: replacement of A by aAb

Chomsky-Schützenberger Hierarchy: Classification of formal languages into four types

(Mathematical) Logic: theory of propositions and their logical connections; fundament of mathematics; propositional logic can be defined as a formal language

Computability Theory: examines the question if a function is computable and especially with which (mathematical model of a) machine the computation can be done

REFERENCES

[1] E. Börger. *Computability, complexity, logic.* North Holland, Amsterdam, 1989.

[2] N. Chomsky. Three models for the description of language. *IRE Transactions on Information Theory*, 2(3):113–124, Sep. 1956.

[3] J. E. Hopcroft, R. Motwani, and J. D. Ullman. *Introduction to Automata Theory, Languages, and Computation (3rd Edition).* Addison-Wesley Longman Publishing Co., Inc., USA, 2006.