

---

## A Context-Free grammar for IUPAC names of hydrocarbons for developing a Syntactic Parser using Swi-Prolog

S. Mahesan

*University of Jaffna, Jaffna, Sri Lanka*

Information retrieval is a necessary task in many applications in any field and retrieval of chemical information is one of them. In addition, predicting chemical structures from their name is also a useful application for students and teachers. For processing the chemical names, there needs to be a systematic or algorithmic way of representing the name. The International Union of Pure and Applied Chemistry has set rules to name a structure for a chemical, starting with hydrocarbons. The rules are not in a suitable form for computer processing. There is only a limited amount of literature on such work, but even in them, the grammar definitions are not made explicit or are implicitly embedded within the source code. The aim of this paper is to present a context-free grammar for organic chemicals so as to process their names for information retrieval or for the provision of structural information for 2-dimensional drawing, and to demonstrate how the names can syntactically be parsed by implementing the grammar in a free and open source language *Swi-Prolog* making use of its automatic backtracking feature and declarative nature. The Backus-Naur notation is used to define the grammar. The definitions are set as Prolog rules correspondingly so as to identify the components of the names with the proper tagging, so that one can make use of the tagged information for other processing as necessary.

**Keywords:** context-free-grammar, hydrocarbon, Swi-Prolog, information-retrieval, structure-drawing

mahesans@gmail.com