

Tree Controlled Grammars

Dominika Regéciová

November 25, 2018

Abstract

Tree controlled grammars were firstly introduced by Culik and Maurer, 1977¹. The basic information about tree controlled grammars will be presented, as well as the known properties of them, and a brief summary of the upcoming article will be shown.

Tree controlled grammars (TC grammars for short) are context-free grammars with an additional restriction upon the generated derivation trees. The restraint is in the form of regular control set. For every word x generated by this grammar, there must be a corresponding derivation tree such that each word is created by the concatenation of all symbols at any level (with the exception of the last one) from left to right from the control set.

Formally, a TC grammar is a pair (G, R) , where $G = (N, \Sigma, P, S)$ is a context-free grammar and $R \subseteq (N \cup \Sigma)^*$. A notorious example of TC grammars is following:

$$\begin{aligned} G &= (\{S, A, B, C\}, \{a, b, c\}, P, S) \text{ with} \\ P &= \{S \rightarrow ABC, A \rightarrow aA, A \rightarrow a, B \rightarrow bB, B \rightarrow b, C \rightarrow cC, C \rightarrow c\} \text{ and} \\ R &= \{S, ABC, aAbBcC\} \end{aligned}$$

This generates a very well known language $a^n b^n c^n$ for $n \geq 1$, which is not context-free.

What is very convenient, every type 0 languages can be generated by TC grammars, and for many languages (even languages which are not context-free) generated by TC grammars, there are parsing methods working in the $\mathcal{O}(n^2)$ time. On the other hand, the use of normal forms can be problematic and it has been shown that some normal forms (e.g. Greibach normal form) cannot be obtained for TC grammars in general.

There are many versions of TC grammars with various computational power. As it was mentioned, every type 0 language can be generated by TC grammar (G, R) , where G is a context-free grammar and R is a regular control set. But what happens when the control set is limited to some type of a subset of the regular grammar? Will the class of the generated languages be the same?

Many articles discuss TC grammars in the context of the type 1 languages, where many subsets of regular TC grammars (like star-free grammars) have the same power as TC(CF, Reg) and the class of context-sensitive languages themselves. However, there are not many papers dealing with TC grammars in the context of the type 0 languages.

A special form of TC grammars, where the control set is union-free regular language, thus generated by regular expressions without the operation union, will be presented. The forthcoming article should prove that this limitation of the control set maintains the computational power of the TC grammars, based on the proof² given by Alexander Meduna and Zbyněk Křivka.

¹K. Culik and H. A. Maurer. Tree controlled grammars. Computing, 19:11, 1977.

²A. Meduna, Z. Křivka: Scattered Context Grammars with One Non-Context-Free Production are Computationally Complete. In preparation.