Journal of Automata, Languages and Combinatoric
s ${\bf 28}$ (2023) 4, 249–277 © Institut für Informatik · Justus-Liebig-Universität Giessen

ANTI-CONTEXT-FREE LANGUAGES

Carles Cardó 💿

Departament de Ciències Bàsiques, Universitat Internacional de Catalunya c/ Josep Trueta s/n, Sant Cugat del Vallès, 08195, Catalunya, Spain ccardo@uic.es

ABSTRACT

Context-free languages can be characterized in several ways. This article studies projective linearisations of languages of simple dependency trees, i. e., dependency trees in which a node can govern at most one node with a given syntactic function. We prove that the projective linearisations of local languages of simple dependency trees coincide with the context-free languages.

Simple dependency trees suggest alternative dual notions of locality and projectivity, which permits defining a dual language for each context-free language. We call this new class of languages anti-context-free. These languages are related to some linguistic constructions exhibiting the so-called cross-serial dependencies that were historically important for the development of computational linguistics. We propose that this duality could be a relevant linguistic phenomenon.

 $K\!eywords:$ anti-context-free languages, context-free languages, local tree languages, projectivity, dependency grammar

1. Introduction

1.1. Historical and Linguistic Motivation

From the beginning of the history of computational linguistics, there was a suspicion that, although context-free grammars could describe a large part of natural languages, they were insufficient [10]. After some discussions over the next decades (see [32] for state of the art at that time), it was confirmed that some natural languages present non-context-free constructions [35]. In this respect, let us consider a pair of formal languages for a fixed alphabet Σ :

$$L_{\text{squa}} = \{ a_1^2 \cdots a_n^2 \, | \, a_1, \dots, a_n \in \Sigma, n \ge 0 \}, \tag{1}$$

$$L_{\text{copy}} = \{ xx \mid x \in \Sigma^* \},\tag{2}$$

A portion of the work presented in this paper has been done while the author was at Departament de Ciències de la Computació, Universitat Politècnica de Catalunya. The research was supported by the recognition 2017SGR-856 (MACDA) from AGAUR (Generalitat de Catalunya).

⁰⁰⁰⁰⁻⁰⁰⁰³⁻³⁸³⁶⁻³⁶³⁵