

PROJECTIVITY PROPERTY IN NLP

ANCA DINU¹

In this paper we present a short survey of the projectivity property and its relevance in NLP. We discuss some graph properties that restrict non-projectivity, the pseudo-projectivity property and the linguistic phenomenon behind non-projectivity. All these aspects influence the accuracy and efficiency in practical applications such as parsing. Until very recently, most practical systems for dependency parsing assumed projectivity (at the expense of accuracy), whereas most dependency-based linguistic theories did not (it is well known that there are some syntactic phenomena such as *wh*-movement in English or clitic climbing in Romance, that require non-projective analyses). We summarize some of the latest work that tries to overcome this drawback.

1. Introduction

The notion of projectivity dates back in the early sixties. It was first introduced by (Lecerf, 1960). The condition of projectivity of a rooted tree has been defined in several ways, some of which have been shown by (Marcus, 1965) to be equivalent. Marcus studied further the projectivity property in (Marcus, 1967). Since then, a vast number of papers address the subject (some of them are listed in (Marcus, 1988)) in fields like linguistics (phonology and syntax), mathematics (projectivity is at the origin of a chapter in graph theory, (Nebesky, 1969)) and even poetics (syntactic distortion in poetry, for example (Cosmas, 1986)).

The initial linguistic motivation for introducing projectivity was the observation that a vast majority of syntactical construction have this property, i.e. in a dependency ordered tree, the subordination (which is the reflexive and transitive closure of syntactic dependency relation) of a node b to a node a , implies that all intermediate nodes are subordinated to a . Although most linguistic structures can be represented as projective trees, it was soon noted that projectivity is too strong a constraint for dependency trees: a number of syntactic phenomena (such as *wh*-movement in English or clitic climbing in Romance, diverse types of fronting, extraposed relative clause, etc) require non-projective analyses.

Non-projectivity has been analyzed from two perspectives: graph properties that restrict non-projectivity and linguistic phenomenon behind non-projectivity.

¹ University of Bucharest.