

Modelling Grammatical and Lexical Knowledge: A Declarative Approach

Palmira Marrafa

University of Lisbon, Faculty of Arts
Department of Linguistics and CLG-Computation of Lexical and Grammatical Knowledge
Research Group (Centre of Linguistics)
Palmira.Marrafa@netcabo.pt

Abstract. This paper depicts the fundamentals of a computational grammar able to provide adequate representations for Portuguese simple sentences with several kinds of ambiguities. Besides the description of the architecture of the system proposed and the way it works, the paper focuses on the discussion of the nature of the specifications to encode in order to get a high level of precision. From a linguistic point of view, an endocentric phrase structure approach is adopted. The information is encoded in a DCG-like formalism, implemented in PROLOG.

1 Introduction

Modelling grammatical knowledge entails the specification of a large set of intertwined syntactic and semantic properties of linguistic expressions, which are highly structured and exhibit local and long distance dependencies ruled by several types of constraints.

In view of the complexity of the information to encode, the development of grammars that are suitable enough both for precision and coverage represents a great challenge.

As well-known, precision and coverage are conflicting requirements of natural language modelling, since a more precise grammar tends to be a more constrained grammar while constraints tend to reduce coverage (see [9] for a brief discussion of this trade-off).

Without neglecting coverage, this work is particularly concerned with precision, an essential requirement both for Theoretical Computational Linguistics central aims and for a wide range of applications.

Accordingly, the fragment of grammar presented here is able to rule out ill-formed expressions and inappropriate interpretations and to assign at least one representation to each well-formed expression for the constructions at issue. It covers the basic structure of simple sentences with several types of predication relations.

Such sentences frequently involve syntactic ambiguity, a major problem for computational natural language analysis.

Despite the complexity of the phenomena involved, the grammar has a suitable level of parsimony, since grammatical rules make appeal to the lexical entries which contain fine grained specifications of the syntactic and semantic restrictions imposed by the lexical units.