

Analysis of Sanskrit Compounds: Constituency or Dependency?

Amba Kulkarni
ambakulkarni@uohyd.ac.in
Department of Sanskrit Studies
University of Hyderabad

Abstract

Sanskrit is rich in compound formation. Therefore a sentential parsing is incomplete without analysing the compounds revealing the relations between its components and also the hierarchies between the nested components.

Constituency analysis for a compound appears as a natural choice given the hierarchical structure with binary compound formation, barring the dvandva (co-ordinating) and bahuvrīhi (exo-centric) compounds involving more than two components. The compound analysis, in addition to the hierarchical structure, also involves identification of the type of the compound. Often a compound in isolation is ambiguous, confusion between the tatpuruṣa and bahuvrīhi being the most common. It is the context that helps resolve such ambiguities in certain cases. Hence access to the context, at least to the sentence, is desirable. Use of 'asamartha' compounds is also very frequent where the expectancies of the modifier or the iic(in initio compositi) of the compound are satisfied by other words in the sentence. This motivates us to extend the dependency sentential parser based on the Pāṇinian framework to compound analysis as well.

In this talk we discuss the challenges in the development of a compound analyser, followed by a computational model for parsing compounds following the dependency framework inspired from the Pāṇinian grammatical formalism and the Indian theories of śābdabodha. We argue how the dependency framework provides an efficient model for compound analysis facilitating its smooth integration with the existing sentential parser. Further it is natural for a human being to cognise the meaning of a compound and also understand the compound structure as a hierarchical binary tree rather than a dependency tree. Therefore we also represent the dependency parse of a compound as a constituency tree thereby producing a human friendly output.