Mathematical modeling of Ākāṅkṣā and Sannidhi for parsing Sanskrit

Amba Kulkarni, Devanand Shukl, Sheetal Pokar

Department of Sanskrit Studies, University of Hyderabad, India,
`apksh@uohyd.ernet.in`

Indian grammatical tradition has provided models for both generation as well as analysis of a language. The Aṣṭādhyāyī provides a framework for generation while the Śābdabodha theories deal with the analysis aspect of a language.

Analysis of a linguistic utterence for understanding involves a process which is an inverse process of generation. The process of generation is deterministic, typically resulting into a single string. The analysis on the other hand may involve non-determinism. To give an illustration, the two sūtras *anabhihite* (A 2.3.1) and *kartṛkaraṇayos tṛtīyā* (A 2.3.18) assign third case to both the kartṛ as well as karaṇa in passive voice. Now, given a sentence in passive construction, one can not decide uniquely the kāraka role of a word in the third case only on the basis of the word-form. One needs to appeal to the 'padārthas' involved and even sometimes to the context.

This problem was well recognised by the naiyāyikas who proposed four conditions viz. ākāṅkṣā (expectancy), yogyatā (mutual compatibility), sannidhi (proximity) and tātparya (intention of the speaker) as necessary conditions for proper verbal cognition.

In this paper we discuss an implementation of a mechanical parser for Sanskrit. We propose a mathematical model to handle sannidhi in order to prioritize the most likely parsed output. For example, in a sentence *saḥ grāmaṁ gatvā prāsādaṁ paśyati*, the parser produces two possible parses: the first one with the grāmaṁ as the karma of going activity and prāsāda as the karma of the seeing activity and the other one with grāmaṁ as the karma of the seeing activity and prāsāda as the karma of going activity. We provide a mathematical criterion to measure the proximity so that the less probable solution is pushed down and the more probable solutions are displayed at the top.

Ākāṅkṣā plays an important role in establishing relations between the words which demand certain expectancies such as verbs, and the words which fulfil these demands such as substantives. These ākāṅkṣās are of different types. Verbs and certain indeclinables demand substantives with certain vibhaktis, the relative pronouns have an expectancy of their matching counterpronouns, and so on. Further these expectancies may be of different intensities. We discuss the implementation of ākāṅkṣā that handles all these various cases.

Finally we also evaluate the resulting parser that handles Ākāṅkṣā and sannidhi and discuss the problems in implementing yogyatā.