

Theories of Śābdabodha and Principles of Information Coding: A computational Perspective

Amba Kulkarni

Abstract

Indians have developed the theories of Śābdabodha to decipher and understand the information coded in language strings. Indian grammatical tradition discusses four factors viz. Ākāṅkṣā (expectancy), yogyatā (mutual congruity), sannidhi (proximity) and tātparya (purport) as essential factors in the process of verbal cognition (Śābdabodha).

These four factors are useful for a computational linguist in understanding the dynamics of information coding in a language string. They provide answers for a) 'where' does a language string code information?, b) 'what' kind of information does it code?, and c) 'how' is such information coded?

In this paper we unfold these factors and explain how do they help in understanding the dynamics of information coding in a language string. Finally we also discuss how these factors also suggest an upper limit on automatic information extraction and the need for BIG DATA in automatic processing of Natural languages.