

Florencia Leonardi

(IME-USP, São Paulo)

Title: The Smallest Maximizer Criterion for Context Tree Selection

Abstract: I will present a constant free procedure for context tree selection called the smallest maximizer criterion. This method also provides, as a by-product, a solution for the problem of optimal choice of the penalty constant when using the Bayesian Information Criterion (BIC) to select a variable length Markov chain. I will also present a consistency result for this criterion and a simulation study comparing our approach with both the standard BIC selection and the Peres–Shields order estimator. I will also show an application of the smallest maximizer criterion to a dataset of codified written texts of Brazilian and European Portuguese to retrieve fingerprints of rhythm. This method assigns different context-tree models to the two dialects of Portuguese. The features of the selected models are compatible with current conjectures discussed in the linguistic literature.