XXIII Congresso de Iniciação Científica da Unicamp 2015
17 a 19 de Novembro
Campinas Brasil

Algorithmic Complexity and Information Theory

Patrick F. Coutinho (IC), Romis R. de F. Attux (PQ)

Abstract

In this work, we performed a comprehensive study about the foundations of *algorithmic complexity* and of *information theory*. Initially, we investigated in detail the first, which deals with discrete sequences and is based on the concept of *Kolmogorov – Chaitin complexity*. Then, we explored key notions of Shannon's information theory, with a special interest in the concept of *entropy*. Afterward, we confronted the two theories, discussing connections between them that are related to coding and compressibility. In the end, we addressed the *blind source separation problem* and proposed an approach to it based on complexity measures, which were estimated by means of compression of the *recurrence plots* of the mixed signals.

Key words: information theory, algorithmic complexity, blind source separation

Introduction

The concept of entropy is associated with the minimal quantity of information needed to losslessly represent a source of information, and can be seen as a bound in terms of data compression[2,4]. On the other hand, the algorithmic complexity of a sequence concerns the minimal length of the program that, being it given to a universal machine with no other input, produces that very sequence[1]. The algorithms of data coding and compression, taken as computer programs, can be used to estimate the algorithmic complexity of a sequence.

These ideas, applied in the field of signal processing, show that the blind source separation (BSS) for linear and invertible mixtures can be solved based on the complexity measures of the mixed signals[4]. Starting from the construction of the recurrence plots[5] of these signals, we adopt the compressed file size of the plots as the estimated complexity.

Results and Discussion

The most common approach to the BSS is the *independent component analysis* (ICA). Nevertheless, as expressed by Pajunen[3] and Soriano et al.[5], this method can be limited in some contexts. The presented approach, which employs the compression ratio of the recurrence plot as the complexity measure of the mixed signals, aims to overcome such restrictions and was verified for linear and invertible mixtures of deterministic or chaotic sources with stochastic ones.

Image 1 shows the result obtained from a recovery of a chaotic source from a mixture with a stochastic one. The recovered source's signal is only reversed with respect to the original source.

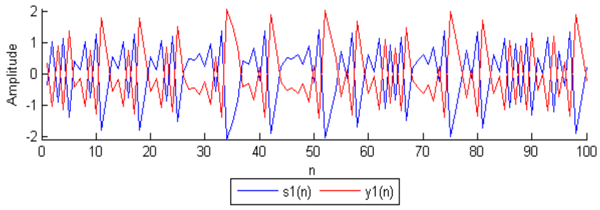

Image 1. Recovering a chaotic source from a linear and invertible mixture with a stochastic source. The color red represents the recovered signal and the color blue represents the original one.

Conclusions

In this work, the foundations of the information theory and of the algorithmic information theory, after they had been studied by the student, were applied in the field of signal processing. The results obtained in several different contexts are significant. The perspective of comparing the results of this methodology with the results of the ICA approach is open.

Acknowledgement

The authors thank CNPq and PIBIC for the financial support.

[1] Chaitin, G. J. *Thinking About Gödel and Turing: Essays on Complexity, 1970-2007*; World Scientific Publishing Company, 2007.

[2] Cover, T. M.; Thomas, J. A. *Elements of Information Theory*, 2nd ed.; Wiley, 2006.

[3] Pajunen, P. Blind Source Separation Using Algorithmic Information Theory. *Neurocomputing*. **1998**, *22*, 35-48.

[4] Shannon, Claude E. A Mathematical Theory of Communication. *Bell System Technical Journal*. **1948**, *27*, 379-423, 623-656.

[5] Soriano, D. C.; Suyama, R.; Attux, R. Blind Extraction of Chaotic Sources from Mixtures with Stochastic Signals Based on Recurrence Quantification Analysis. Digital Signal Processing. **2011**, *21*, 417-426.