Physica A 422 (2015) 66–72

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

A measurement of disorder in binary sequences

Longyan Gong [a,b,*], Haihong Wang [a,b], Weiwen Cheng [b], Shengmei Zhao [b]

[a] *Information Physics Research Center and Department of Applied Physics, Nanjing University of Posts and Telecommunications, Nanjing, 210003, China*

[b] *Institute of Signal Processing and Transmission, Nanjing University of Posts and Telecommunications, Nanjing, 210003, China*

HIGHLIGHTS

- A complex quantity, A_L, is proposed for binary symbolic sequences.
- It is well consistent with various entropies.
- It is a useful quantitative measure of disorder.
- It has many advantages.
- To existing measurements, it is a very nice complement.

ARTICLE INFO

Article history:
Received 27 August 2014
Received in revised form 31 October 2014
Available online 9 December 2014

Keywords:
The degree of disorder
Binary sequences
Time series analysis

ABSTRACT

We propose a complex quantity, A_L, to characterize the degree of disorder of L-length binary symbolic sequences. As examples, we respectively apply it to typical random and deterministic sequences. One kind of random sequences is generated from a periodic binary sequence and the other is generated from the logistic map. The deterministic sequences are the Fibonacci and Thue–Morse sequences. In these analyzed sequences, we find that the modulus of A_L, denoted by $|A_L|$, is a (statistically) equivalent quantity to the Boltzmann entropy, the metric entropy, the conditional block entropy and/or other quantities, so it is a useful quantitative measure of disorder. It can be as a fruitful index to discern which sequence is more disordered. Moreover, there is one and only one value of $|A_L|$ for the overall disorder characteristics. It needs extremely low computational costs. It can be easily experimentally realized. From all these mentioned, we believe that the proposed measure of disorder is a valuable complement to existing ones in symbolic sequences.

1. Introduction

Symbolic sequences have been under study for a long time in many areas, including physics [1,2], biology [3–5], information science [6], economics [7], and linguistics [8,9]. A fundamental question about the subject is how disordered a certain sequence is [10]. A similar one is that which sequence is more disordered for two different sequences [11,12]. Unfortunately, even today there is no one definite answer. Quantitative studies require appropriate measures of the degree of disorder. In researches, different measures (methods) have been proposed, in a sense, they are complementary to each other. For example, the methods include the autocorrelation function analysis, the Fourier power spectrum analysis, the

* Corresponding author at: Information Physics Research Center and Department of Applied Physics, Nanjing University of Posts and Telecommunications, Nanjing, 210003, China.

E-mail addresses: lygong@njupt.edu.cn (L.Y. Gong), happy.haihong@163.com (H.H. Wang), wwcheng@njupt.edu.cn (W.W. Cheng), smzhao@njupt.edu.cn (S.M. Zhao).

http://dx.doi.org/10.1016/j.physa.2014.11.059
0378-4371/

detrended fluctuation analysis, the duration time analysis, the entropy analysis and others (see Refs. [10,13] and references therein). The measures, for instance in aperiodic sequences, can be based on an information theoretic measure [11,12], the diffraction pattern of X-ray [1,14], and the properties of the electronic energy spectrum of a one-dimensional tight-binding model of lattices [1,15]. As they are from different points of view, even for a same sequence they may produce inconsistent estimators about the degree of disorder [12,10,16,17]. Therefore, it is worthwhile to develop techniques to understand the disorder in sequences.

The simplest symbolic sequences are binary ones, which only include two symbols. Many signals are binary variables, or can be reduced to binary sequences. Even for the elementary case, a generally accepted measure of the degree of disorder is lacking. In this letter, we will propose a complex quantity, A_L, to measure the disorder of binary sequences. As examples, we respectively apply A_L to two kinds of typical random binary sequences and two typical deterministic quasiperiodic binary sequences. With its help, we compare the disorder of the Fibonacci sequences to that of the Thue–Morse ones. We find that the modulus $|A_L|$ is well consistent with various entropies (the Boltzmann entropy, the metric entropy, or the conditional block entropy) and other quantities in the analyzed sequences, so it is a novel and useful quantitative measure of the degree of disorder. In addition, it has many advantages. To existing measurements, it is a very nice complement.

2. The measure of disorder

Let us start with the definition of the measure of disorder for binary numerical sequences. We consider a numerical sequence

$$S_L = \{\varepsilon_1, \dots, \varepsilon_\ell, \dots, \varepsilon_L\}, \tag{1}$$

where the subscript ℓ plays the role of discrete time or space, $\varepsilon_\ell = 0$ or 1, and L is the sequence length. We propose a complex quantity

$$A_L = \frac{1}{\sqrt{M}} \sum_{\ell=1}^{L} \varepsilon_\ell \exp(\mathrm{i}2\pi\ell/L), \tag{2}$$

where $M = \sum_{\ell=1}^{L} \varepsilon_\ell$ and $i = \sqrt{-1}$.

Inspired by two aspects, we introduce the A_L. One comes from the many-body localization theory in condensed matter physics [18–20], and the other relates to the diffraction theory for complex gratings [21]. In one aspect, Resta has used a quantum-mechanical position operator and defined a fundamental complex quantity, $\mathbb{Z}_L$, to study the localization properties of many-body extended quantum systems within periodic boundary conditions [18]. For the simplest case that only a single electron moving in a one-dimension lattice with L sites,

$$\mathbb{Z}_L = \sum_{\ell=1}^{L} |\psi_\ell|^2 \exp(\mathrm{i}2\pi\ell/L) \tag{3}$$

with $\sum_{\ell=1}^{L} |\psi_\ell|^2 = 1$, where ψ_ℓ is the probability amplitude of the electron at the ℓth site. Eq. (3) shows that for an extreme delocalized state that $\psi_\ell = \frac{1}{\sqrt{L}}$ for all ℓ, the modulus $|\mathbb{Z}_L| = 0$, and for an extreme localized state that $\psi_\ell = \delta_{\ell\ell_0}$, $|\mathbb{Z}_L| = 1$, where ℓ_0 is a given site and δ is the Kronecker delta function. It has been found that $|\mathbb{Z}_L|$ and related quantities provide an appropriate measure of localization properties of wavefunctions [18–20]. Here we borrow it to study the degree of disorder of the binary numerical sequence S_L. For this, we suppose that

$$\psi_\ell = \begin{cases} 1/\sqrt{M} & \text{if } \varepsilon_\ell = 1, \\ 0 & \text{if } \varepsilon_\ell = 0. \end{cases} \tag{4}$$

It describes the special case that the electron uniformly occupies at sites with $\varepsilon_\ell = 1$, and there is no electron at sites with $\varepsilon_\ell = 0$. In physics, such wave functions can be realized at the condition [22] that the on-site potential is finite at sites with $\varepsilon_\ell = 1$, while it is positive infinity at sites with $\varepsilon_\ell = 0$. And there are enough hopping links [23] between sites with $\varepsilon_\ell = 1$. Using Eqs. (3) and (4), we obtain the numerical relation that $A_L = \sqrt{M}\mathbb{Z}_L$. In other words, our proposed A_L is a rescaled quantity of $\mathbb{Z}_L$ at the specific case.

In the other aspect, the Fraunhofer diffraction about complex gratings with very narrow slits is considered [21]. For the normal plane-wave incidence, the interference wave can be written as

$$\mathbb{A}_L = \sum_{\ell=1}^{L} a_\ell \exp\left(\mathrm{i}\frac{2\pi\ell d}{\lambda}\sin\theta\right), \tag{5}$$

where a_ℓ is the amplitude of diffraction wave with wavelength λ from the ℓth slit, d is the distance between two nearest slits, θ is the diffraction angle. To get Eq. (2), we use the relation that $\sum_{\ell=1}^{L} |a_\ell|^2 = 1$, which represents all the energy of the transmission light is set to unit one. We first set

$$a_\ell = \begin{cases} 1/\sqrt{M} & \text{if } \varepsilon_\ell = 1, \\ 0 & \text{if } \varepsilon_\ell = 0. \end{cases} \tag{6}$$