A Biological Model with Substitution Operator

Ludmila de Pinho Cavalcanti[1], *Dr. Alex Dias Ramos*[1] , *Dra. Caliteia Santana de Sousa*[1]
[1] Universidade Federal de Pernambuco

Abbreviated abstract: During DNA replication, several variations are likely to occur. That is, during the process of synthesizing a new strand, alterations may occur in the genetic code (such as insertion, inversion, duplication, deletion, …) that trigger poor cell development. The idea developed in this work, thus far, is the study of a model that analyzes the behavior of finite sequences of two elements (zero and one), that is, words.

VIRTUS IMPAVIDA

ludmila.cavalcanti@ufpe.br

POM
Photonics Online Meetup

Motivation and some Definitions.

Based on DNA replication, Li (1989) proposed a model known as the expansion-modification system, which works as an approximation of the process of updating the DNA sequence. This model has two behavior, one of which preserves the long-range correlation, while the other tends to destroy it.

1. We define a **word** as any finite sequence of zeros and ones. And define the **size of a word** as the quantity of zeros and ones that compose the word.

2. The set of all possible words is defined as **dictionary** denoted by dic.

3. Let $(X_n) \in \chi_\Omega$, where χ_Ω is the set of random words. This sequence **converges** for a measure $\mu \in \mathcal{M}$, if for all word W ∈ dic we have:

$$\forall\, \varepsilon > 0, \exists\, n_0 \in \mathbb{N}\, ; n > n_0 \Rightarrow \left|d_{X_n}(W) - \mu(W)\right| < \varepsilon.$$

4. An application P is **consistent** if for any $\mu \in \mathcal{M}$ and sequence $X_1, X_2, X_3, \ldots \in \chi_\Omega$ such that $X_n \to \mu$, the sequence $X_1P, X_2P, X_3P, \ldots$ converges and is unique for any $X_n \to \mu$.

5. The operator P is defined as **ergodic** if for every measure $\mu \in \mathcal{M}$ and for every (X_n) where $X_n \to \mu$, we have $\lim_{t\to\infty} X_n P^t = \mu$.

ludmila.cavalcanti@ufpe.br

Photonics Online Meetup

Model

The model P, a particular case of Li Model, is given by

$$P: \begin{cases} 0 \rightarrow \begin{cases} 1, & \text{with probability } p. \\ 00, & \text{with probability } 1-p. \end{cases} \\ 1 \rightarrow 1 \end{cases}$$

Suppose we have a sequence that has two symbols, one (1) and zero (0), and at each instant of time t, the symbol zero can suffer a change based on the model P above, which will be a modification, that is, the symbol will change to 1 with probability p at the next time t+1. Alternatively, it can change to an expansion, that is, expand to a duplicate of yourself (00) with probability 1-p at instant t+1. And the symbol one will remain the same. Fig. 1 describes the evolution of a particular word.

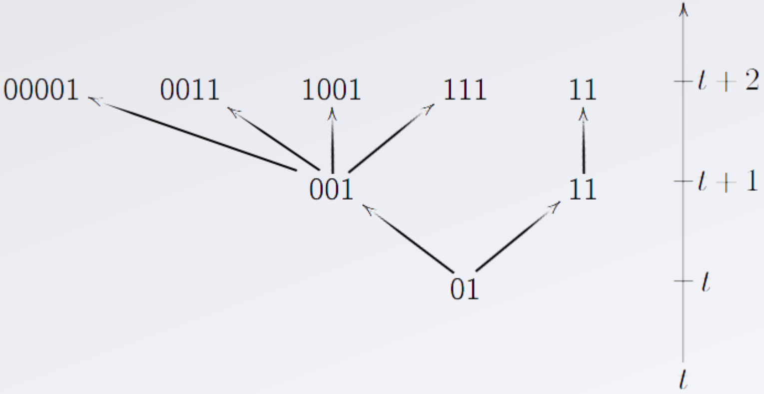

Fig 1: Evolution of the word W = 01 at each instant of time t.

ludmila.cavalcanti@ufpe.br

Results

Theorem 1: The operator P is consistent.

Theorem 2: Let p be the probability of the model:

a) If p = 1, then $\mu P^t = \delta_{1_Z}, \forall \mu \in M$

b) If p = 0, then: $\lim_{t \to \infty} \mu P^t = \begin{cases} \delta_{0_Z}, se\ \mu(0) > 0. \\ \delta_{1_Z}, se\ \mu(0) = 0. \end{cases}$

Theorem 3: Let p ∈ $(0,1)$. In our process, we have only one invariant measure if p ≥ $\frac{1}{2}$ and there are at least two invariant measures if $p < \frac{1}{2}$.

Theorem 4: Let p be the probability of the model. For $p \geq \frac{1}{2}$, P is ergodic.

References

1. LI, W. Generation non-trivial long-range correlations and 1/f spectra by replication and mutation. International Journal of Bifurcation and Chaos in Applied Sciences and Engineering, World Scientific, v. 02, n. 01, p. 137–154, 1992.
2. TOOM, A.; RAMOS, A.; ROCHA, A. V.; SIMAS, A. Random Processes with Variable Length. In: 28-o Colóquio Brasileiro de Matemática. [S.l.]: IMPA, 2011. ISBN: 978852443316.

ludmila.cavalcanti@ufpe.br