

Mathematical Structures Applied in the Characterization of the Genetic Code

Anderson Oliveira

Federal University of Alfenas

Mathematical modeling of the genetic code enables, among other aspects, the analysis, interpretation, and characterization of properties associated with amino acids and possible interferences in various situations, such as genetic mutations. Several structures can be used in this process, such as the so-called Hasse diagrams, presented by Fernandes and Oliveira (2021). In addition, through the Gray code and the Boolean hypercube, several properties and characteristics can be analyzed, as shown by Jimenez-Montano and La Mora-Basanez (2002) and Ribeiro and Oliveira (2024). The genetic code represents the association of triplets found in messenger RNA, formed by nitrogenous bases, and the amino acids that are in the structures of proteins. The triplets of nitrogenous bases, mentioned above, are called codons. There are 64 possible triplets which correspond to 20 amino acids, which are the basic units that form proteins. The correspondence of a given amino acid with a codon occurs through the genetic code, as presented in Alberts (2017). Faria (2011) presented a modeling of the DNA structure, through a mapping of the genetic code alphabet, which is related to the set of nitrogenous bases $N = \{A, C, G, T/U\}$, (A - adenine, C - cytosine, G - guanine and T/U - thymine/uracil) and the 4-ary alphabet of the ring structure, denoted by $\mathbb{Z}_4 = \{0, 1, 2, 3\}$. The associations of the biological alphabet with the mathematical alphabet allow the identification of 24 permutations, which are organized into three labels, which aim to determine the best association of each of the symbols of the set N with the corresponding symbols of the set $\mathbb{Z}_4$ and the opposite. Each permutation has its own characteristics, allowing this mapping to be organized into three sets, called labels A, B and C, each one containing 8 permutations, according to the geometric characterization associated with the permutations, as per the construction presented in Faria

(2011). Another important contribution to the process of modeling and analyzing the genetic code is presented in Hage and Harary (2002), through a Boolean hypercube and its relationship with a binary code, showing constructions for dimensions 1 to 4 of the hypercube. In Jimenez-Montano, La Mora-Basanez and Poeschel (2002) a representation of the genetic code is presented using a 6-dimensional Boolean hypercube, constructed from the Gray code table, in which the vertices of each cube are represented by the codons of the genetic code. Based on this work, Ribeiro and Oliveira (2024) provided a systematic procedure for the construction and analysis of Boolean hypercubes associated with each of the labels of the genetic code, where several analyses and interpretations of the constructions made are performed. The hypercube simultaneously represents the entire set of codons and keeps track of which codons are adjacent. The construction of the Boolean hypercube is performed from all the connections, made by identifying the difference of a bit existing from one vertex to the other, from dimension 1 to dimension 6. Based on the information presented previously and important questions still open about the mathematical modeling of the genetic code, the main of this work is to present and to reflect on some contributions using mathematical structures in the process of characterizing the genetic code, through the construction and analysis of these structures, seeking to motivate future research in the area. **Keywords:** Biology. Gray Codes. Hasse Diagrams. Boolean Hypercubes. Algebra. Geometry. **Acknowledgments:** Federal University of Alfenas and the Postgraduate Program in Applied Statistics and Biometrics. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001

In Special Session 47 on Mathematical Tools with Applications in Quantum and Genetic Coding at the Mathematical Congress of the Americas 2025.