

Information theory, error correcting codes, and the living world

G rard Battail

E.N.S.T., Paris, France (retired)

gbattail@club-internet.fr

The conservation of genetic information through the ages can not be explained unless one assumes the existence of genomic error-correcting codes, as we did in earlier works. Shielding by phenotypic membranes does not protect the genomes against radiations. Only means intrinsic to the genome itself can do so. Error-correcting codes are by-products of information theory. They achieve reliable communication over unreliable channels, so paradoxical it may look. They are among the most powerful tools available to communication engineers, whose experience constitutes a strong experimental proof of their efficiency. As a subsidiary hypothesis, we assumed moreover they take the form of 'nested codes', i.e., that several codes are combined into a layered structure which results in an unequal protection: the older and more fundamental parts of the genomic information are better protected than more recent and less essential ones.

Based on these hypotheses, fundamental results of information theory explain basic features of the living world, especially the discreteness of species and the possibility of a hierarchical taxonomy. They account for some yet unexplained facts, e.g., that life proceeds by successive generations, as well as the trend of evolution towards complexity. Also, the saltationist vision of evolution is strongly supported. Trying to identify the hypothesized genomic error-correcting codes, we were led to the further hypothesis that the many physical, steric, chemical and linguistic constraints to which the DNA molecule and the proteins for which they code are submitted act as 'soft codes' which provide the necessary error-correcting ability. It turns out that we then obtain a picture of the living world which closely resembles the 'organic codes' described by Barbieri although he arrived at this concept by completely different ways.

This paper is intended to some remarks of methodological or epistemological character on the way information theory and error-correcting codes can help understanding the living world. That they explain biological facts left unexplained by today's biology seems to validate our hypotheses, but the direct experimental identification of genomic error-correcting codes still lacks. It would obviously require the active collaboration of practicing genetists.