Review of: "Modeling the structure and evolution of cultural information as Quasispecies"

Hugo I. Cruz-Rosas[1]

1 Universidad Nacional Autónoma de México

Potential competing interests: No potential competing interests to declare.

This manuscript presents a model that shows a link between biological evolution and the evolution of information transmitted by language systems in modern human culture.

The model allows for identifying the role of copying errors in information encrypted in texts or symbolic systems (such as partiture music), driving to generate new symbol chains and increasing the diversity of "meaning" (messages) in a population. In this sense, an optimum interval for error rate makes possible the emergence of sequences that can be copied faster in transmitting information (Figure 1).

An interesting analogy is depicted between the convergent and divergent evolution linking them with the convergence and divergence of the information content carried in chains of symbols. Also, the conclusion that machines and AI tend to dominate the management of information is supported by the model based on the incapability of individuals to process the amount of current knowledge and the increasing rate of diversifying information.

However, although the author works with "Cultural Information," the model only considers a particular kind of that: those codified (encrypted) in chains of symbols (like a string). In the same manner, Dr. Stevenson regards genetic information, leaving out the phenomena where information is stored and transmitted without ordered sequences (of monomers). The cultural information, as "cultural," includes non-sequences-based systems such as beliefs, traditions, and no lenguaje-based art. Likewise, biological information includes processes non-based on molecular strings, such as electric and chemical synapses, cellular communication employing hormones or feromones, and the processes of transduction of environmental stimuli, where strings do not transmit information (information is contained in the molecular geometry or the amplitude of membrane potential changes). Reducing "cultural" to oral or written information (cryptic information) and also reducing biological information to genetic one is an oversimplification of these phenomena.

The use of the terms: "cultural information (and "evolution")," "human communication," "communicated information," "language (evolution)," "written language," and "sentences" seems to be employed as synonymous; reinforcing the oversimplification of "culture" to "language" placing all the above terms on the same conceptual plane. The same applies to using "biological evolution" or "biological information," but it is only considered the nucleic acids (RNA) in the model.

Additionally, there needs to be a clearer presentation of the considered sequences. At some points in the introduction, there are references to sequences of letters (like strings), although the sequences that enter into the model are those composed of words (a set of letters with a meaning). The above has implications for the model because a string like "t-h-

e- -c-a-t- -i-s- -b-l-a-c-k" can lose its significance by permuting or replacing some symbol in its sequence (ej. "t-h-e- -t-a-c- -i-s- -b-l-a-c-k"). Meanwhile, another sequence like "the- -cat- -is- -black" can hold its meaning notwithstanding the error in copying its sequence (due to each position being a word in the "string"). In the first sections of the manuscript, there needs to be more clarity about these two kinds of strings. In this sense, the results of the present preprint only are valid for sequences composed of words (a set of symbols with intrinsic meaning). Only in this context the assumption that can emerge sequences with a major rate for copying is valid for containing information.

On the other hand, the modern paradigm explaining biological evolution incorporates (natural) selection as a mechanism acting over variation (from random genetic mutations), causing differential survivance between alternative versions of individuals in a population. Only considering variations (errors) without selection is also an oversimplification of biological evolution.

In my opinion, the presented model is valid only to show the analogy in the evolution of information transmitted by language (based on sequences of symbolic tokens) with the evolution of sequences of nucleotides varying randomly (similar to a process of genetic drift).

I suggest that the author clarify the used "cultural information" (most of all, the "cultural" component) and justify why language evolution is extensive to whole culture evolution.

Finally, a couple of errors must be attended to:
1) When equations 2a and 2b are presented, the author miswrote "equations 5a, and its linear equivalent, 5b..."
2) When equation 7 is presented, the author miswrote "Substituting equation 4 into equation 6 gives us, " but the substitution was done with equation 5 into equation 6.