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

Milind Watve

Potential competing interests: No potential competing interests to declare.

What insights analogy can contribute?

The paper by David Stevenson “Modeling the structure and evolution of cultural information as Quasispecies” is based on the analogy between biological and cultural evolution. Much part of the analogy by itself is not new, as pointed out by some of the earlier reviews. The author has tried a formal model to extend the analogy into something that should add deeper insights into language and cultural evolution.

Analogies are useful in science but they themselves do not make science. The important question is what insights you obtain from the analogy that were not possible without the analogy.

The same can be said about models. Models themselves do not make science. What insights or useful predictions a model can give, that could not be given without the model, is the critical question. For this to happen, a paper should spell out the purpose of building the model, or the questions whose answers are sought, prior to describing the model and later highlight the insights obtained from the modeling exercise. Unfortunately the reader does not find both these components clearly in the paper.

The manuscript may not contribute something new and substantial, but it does trigger thinking. The strength of the Qeios model of open peer review lies in this. The discussion in the public domain may have seeds of novel lines of thinking that can stimulate further work. Keeping the comments confidential and ending up in a dichotomous accept-reject decision is too dumb a method that suppresses potential seeds of new thinking. Open peer reviews is the right kind of fertile land in which seeds of novel thinking can germinate.

With the conventional system of peer review, I would have recommended rejection of this manuscript, although I must admit that it did trigger my thinking on some interesting issues and that is important. At the moment my thinking is not worth writing another full paper, but at the same time, not so worthless so as to suppress it completely. It is possible that it triggers more work, by the authors of this manuscript or by someone else, including myself at some stage. Therefore for the progress of science it is necessary to adopt an open peer review system and terminate the dumb and underproductive accept-reject system.

The interesting possibilities raised by the paper do not directly flow from the model. The interesting part of the discussion could have been communicated equally well even without the model. Since the purpose for the model was not clear to me from the MS as written, I did not go into the mathematical details of the model and refrain from commenting on it.

The important questions and ideas raised by reading the paper are,

1. The question that needs to be addressed with priority is whether cultural evolution is mutation limited or selection limited. The paper relies substantially on error/mutation rates with a hidden assumption that it is mutation limited, but no evidence for that is given. It should be possible to address this question using social media data but that needs looking beyond the language errors alone.
2. There are grossly two major levels of selection working on language and communication. One is selection acting on the form and structure of language. For example social media strongly select for shorter length and therefore new short phrases, acronyms, emoji and such forms would rapidly get selected. This is easier to model and analyze from empirical data, which the author has touched upon. Fitness in this case can be quantified easily.

The other level of selection, which is more important for cultural evolution, is the selection acting on the perceived meaning and social relevance of a message. This is more difficult to quantify as well as model. Currently the paper does not perceive the two levels as separate and incorporate them in the model. This is not an easy task but needs substantially more thoughtful inputs. It is possible that social media are also influencing the second level of selection. They might be selecting for more dichotomous and polarizing contents. Quantifying this is certainly a greater challenge.

The quantifiable parameters used by the author such as sentence length, hamming distance and Shannon diversity are relevant to level 1 but not necessarily to level 2. This limitation needs to be at least clearly acknowledged.

1. While using any analogy it is necessary to appreciate the differences along with similarities. One major difference between biological and cultural evolution is that language has reverse flow of information which biological inheritance does not have. In genetic information although selection may act on the phenotype, what ultimately gets selected is the sequence information. In language communication, one can grasp the meaning and forget the words and then may reconstruct the same or slightly modified meaning in completely different words. Here the meaning gets selected and not the sequence of words. This is not at all analogous to genetic information. Since this is common in human communication, the model required for selection level 2 above needs to be fundamentally different and that is not considered by the author in his model.
2. The manuscript has substantial emphasis on mutation rates. But in human perception of language there is a strong counteracting force. We tend to subconsciously correct and read words. This phenomenon is demonstrably very frequent and therefore the actual mutation rate relevant to evolution can be substantially lower than what is estimated by analysis of text errors. Again without answering whether cultural evolution is mutation limited or selection limited, the relevance of error rate measurement to rate of evolution will not be clear.
3. Although the analogy with mutation and selection is fair, I don't understand the benefit of quasispecies analogy. Species is a difficult concept in biology itself, but apart from that how a quasispecies adds to our understanding of language and cultural inheritance cannot be appreciated from the paper as written currently. Greater clarity is needed on this point, if the analogy is to be really useful.

Our understanding of evolution of human communication and concept building is yet very primitive. Analogy with biological

evolution can certainly be useful but with its own limitations as well. But the analogy by itself is not new and the field needs to progress further with more insightful contributions, building ideas and models and refining our understanding of the evolution much more complex than biological evolution.