Review of: "Self-Replication, Spontaneous Mutations, and Exponential Genetic Drift in Neural Cellular Automata"

Bradly Alicea

Potential competing interests: No potential competing interests to declare.

This is a good paper with some interesting results. Your interactive notebooks are also quite interesting. I have some technical questions, particularly in terms of implementing the evolutionary aspects of the simulation. Specific comments by section below.

Abstract:

the word "inheritable" should be "heritable" (both here and throughout the manuscript).

Introduction:

Is a closed-world approach consistent with Open-ended Evolution? It is not clear in the first paragraph. You need to resolve the first sentence with the canonical definition of OEE.

The "laws of biology" sentence is also unclear. Is this with respect to Cellular Automata? Also, clarifying the specific laws you are referring to (e.g. forces of evolution vs. other laws of physiology, growth, or ecology).

The "one-to-one mapping" comment in the penultimate paragraph of this section is key. In fact, it deserves to be the beginning of its own paragraph or perhaps as the beginning of a subsection.

Question: can an NCA be as flexible as you assume? What is the consequence of this, assuming you have created patterns (which you have)? It seems that there is little distinction required between individuals and a variety of individuals (populations).

Methods and Results:

While your definition of genetic drift is not unacceptable, it also requires a way to rule out the potential functional aspect of a mutation. For example, non-neutral mutations might do more than prevent replications with a single change. In your definition, it seems that drift would necessarily correlate to mutation rate, so that it simply reflects changes in mutation rate and not mutational load.

Choosing the first offspring can also bias your measure of genetic drift -- as you are subsampling one set of mutational pathways at random (stochastic selection is often not equivocal). Example of sampling bias from bacterial colony selection:

Mahilkar, A., Raj, N., Kemkar, S., and Saini, S. (2022). Selection in a growing colony biases results of mutation

accumulation experiments. Scientific Reports, 12, 15470.

Genomic MSE between ancestor and descendent is the definition of evolutionary divergence, correct? Why is there a dip in the divergence function in Figure 1 (right-hand side graph)? Is this an effect of stochastic sampling, movement along the fitness landscape, or a consequence of something else?

Does the transplantation process of preventing pattern crowding also introduce (and favor) genetic drift? This is similar to the subsampling bias mentioned earlier.

I almost think that this work (does not need to be in this paper) would benefit from a benchmark to understand the role of neutrality in more detail. One way to do this is to implement the same genome in a standard GA with a fitness function for reproductive fecundity. How many mutations does it take to shut off reproduction? A host of issues like this could be tested. That way you could test assumptions about neutrality.

If you do not use a fitness function, how do you train a model to achieve the target state? Adding a description of how this process works might be useful to non-NCA experts.

The Phenotype MSE plot in Figure 6 is more what I would expect from an experiment like this (as compared with the plot for Figure 1 and genotype plot in Figure 6). Why the disparity between genotype and phenotype? Same comment for the graphs in Figure 7, although mapping the phenotypes onto that graph might implicitly explain it (not clear to the reader).

There is a type in the paragraph after Figure 6: "tThe" should be "The".

The "paint by numbers" criticism could also be resolved by clarifying the role of training in matching targets with the genotype-phenotype mapping. In other words, how supervised is your training process?

According to my impressions, this is the correct order: produce clones with each seed, each seed is a self-similar pattern, patterns merge and become phenotypic modules. How does this compare with Hox family gene expression during biological pattern formation? From what I understand, Hox genes are highly conserved, experiencing purifying selection rather than accumulating neutral mutations. The former (purifying selection) makes pattern formation that is both spatially distinct and reproducible. How do Neural CAs compare?