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

Kamalika Bhattacharjee

Potential competing interests: No potential competing interests to declare.

Comments:

I am not an expert on neural cellular automata having limited experience with neural network. I accepted to review this work to explore application of neural network on CA and vice versa.

From the perspective of a cellular automaton enthusiast, I have several questions on this paper. First and foremost is, how is the cellular automaton rule defined here? It is only mentioned that, the CA is 2-D and

“The main characteristic of a NCA is that these update rules are encoded by a Neural Network (Noted NN in Fig. 1). In this paper, the state is a vector of 16 real values between -1 and 1, with the first 4 values corresponding to RGBA channels used to render an image on the NCA’s grid.”

Even if the rule is “generated” by the neural network, how does it look like? For each of the cases, is the same CA rule generated? The paper is completely silent about this.

If I consider that by the neural network, a CA rule is generated that can create the patterns, “self-replication, spontaneous mutations, and exponential genetic drift”, then, after getting the rule, simply by evolving it with the same initial configuration, the same behavior should be reproduced. Is the author being able to do so? Please show examples.

Also, what is the lattice size and boundary condition for all cases, is it periodic? And is the neighborhood considered Moore neighborhood for all cases? A CA can not defined without its rule, lattice structure, neighborhood vector and state. So, the author should mention the missing values to give a proper definition of CA.

Other comments:

What do you mean by “world-rule approach of non-neural cellular automata”? What is the size of “world” here?

How does the training work? Can you formally define the training stage? The way it is described is vague.

In the section “Modified training for self-replication”, it is mentioned that “In some experiments we instead use synchronous update rules (all cells are updated simultaneously)”. When is it applied? Show example with both synchronous and asynchronous update over same initial setup to understand the difference.

What is meant by “loss converges”?

In the section “Calculating the genetic drift”, it is mentioned that an egg is “a small, square clump of black pixels”. Then it is mentioned that, “We record the value of the state of all cells in the first egg laid by an organism, and call this value the DNA of the organism.” What do you mean by this? All states are black as per your definition. Then what exactly are you recording? This is very important as the rest of the paper is dependent on this DNA. Please define this formally.

Define “Mean Squared Error (MSE)” with respect to your work.

Formally define “phenotype”. What is “phenotypic drift”?

“In practice, the NCA becomes a replication function for “exactly A”, and any minute deviation A* from the target pattern stops the replication.” -- How do you measure this A*? Not clear.

“In practice, in our small grid, the patterns rapidly crowd each other, so to investigate replication and mutations, we cut out individual patterns and transplant them to an empty grid.” -- If you do that, initial configuration will be changes, so effectively the evolution will be changed.

“The main source of stochasticity in the NCA is the asynchronous update rule. The synchronous model’s training is more brittle and often fails to converge, especially if the training has several targets.” -- Clearly mention when which rule is applied. Is there any reason for choosing that rule at that time?

There are several other lines which seem vague and rather philosophical than having defined with proper mathematical reasoning. Overall, the paper is not well-written and is difficult to follow. The flow of the paper needs to be rechecked. There is sudden capitalization of letters in the paper without any appropriate reason. For example, “Life”, “Earth” etc. This can be avoided. Also, there are many grammatical errors and colloquial symbols which should be corrected while proof reading.

The paper is a nice way to start working, but, in my opinion, needs to be theorized and formalized with more rigorous experimentation to be a good contribution to the general cellular automata community.