

# Open-Ended Evolution in Cellular Automata Worlds

Peter Andras<sup>1</sup>

<sup>1</sup>School of Computing and Mathematics  
Keele University  
p.andras@keele.ac.uk

## Abstract

Open-ended evolution is a fundamental issue in artificial life research. We consider biological and social systems as a flux of interacting components that transiently participate in interactions with other system components as part of these systems. This approach and the corresponding reasoning suggest that systems able to deliver open-ended evolution must have a representation equivalent of Turing machines. Here we provide an implementation of a such model of evolving systems using a cellular automata world. We analyze the simulated world using a set of metrics based on criteria of open-ended evolution suggested by Bedau et al. We show that the cellular automata world has significantly more evolutionary activity than a corresponding random shadow world. Our work indicates that the proposed cellular automata worlds have the potential to generate open-ended evolution according to the criteria that we have considered.

## Introduction

Open-ended evolution has been a core topic of artificial life research since the beginning (Langton, 1995). The inspiration comes from natural biological and social systems that evolve apparently in a limitless, open-ended manner. Replicating such evolutionary processes is naturally required for developing evolving artificial life that is able to deal with novel problems that were not encountered before. The first systematic and analytic approach to quantify what open-ended evolution means in artificial life context was proposed by Bedau (1992; Bedau et al, 1998). Since then several systems have been built that aimed to provide a platform for developing systems with open-ended evolution, these include the Tierra (Ray, 1991; Shao and Ray, 2010), Geb (Channon and Damper, 1998; Channon, 2006), PolyWorld (Yaeger, 1994; Yaeger and Sporns, 2006), Avida (Ofria and Wilke, 2006; Fortuna et al, 2017), Aevol (Knibbe et al, 2007), Chromaria (Soros and Stanley, 2014) and others. Among these, it has been shown, for example, that the evolutionary process supported by the Geb system satisfies the criteria of open-ended evolution (Channon and Damper, 1998) laid down by Bedau (1992; Bedau et al, 1998) and also a further modified criterion proposed by Channon (2006). Cellular automata based simulations of evolution have been used since the beginning of artificial life research (Langton, 1984). Among these, we note the work of Sayama (1998, 1999), which shows how two dimensional cellular automata worlds can generate new species, although in this case the evolution

converges to the same kind of dominant species in all cases. We also note the recent work of Adams et al (2017) that investigates the use of cellular automata for open-ended evolution.

Biological and social systems undergo long-term evolution and lead to a large variety of evolved adapted systems. These systems exist in the context of variable environment and they can be seen as a flux of components that originate from the environment, end up in the environment, and transiently participate in interactions with other system components as part of the system (Alberts et al, 2008; Luhmann, 1996). This suggests that open-ended evolution simulations should be considered in the context of a variable environment, where individuals must pick resources from the environment and discard their waste into their environment. The resources are used to maintain the existence of the individual. The environment for any individual is made of the other individuals and possibly other environmental elements, some of which may be used as resources.

The above view of biological and social systems implies that these systems must compute somehow the way how they should interact with their environment and in an abstract sense they should be able to make predictions about their environment (Andras, 2011). The systems which make better predictions will have better chance to survive, expand and generate descendant systems. The further consequence of this view of such systems is that these systems must be equivalent in some abstract sense with Turing machines and must have correspondingly infinite representations (Andras, 2011), similar to the category theory representations of Turing machines (Pierce, 1991). The implication of this for building models of open-ended evolution is that these models must allow somehow the work with such infinite representations of the systems that develop.

Here we propose a cellular automata (Wolfram, 2002) world in which individuals are bit strings that may survive and reproduce in certain conditions. The proposed cellular automata world can be set up such that it allows the equivalence with Turing machines and provides an implicit way to deal with the corresponding infinite representations of evolving systems. The cellular automata world is analyzed using the metrics and criteria proposed by Bedau (1992; Bedau et al, 1998) and further derived metrics and criteria for the assessment of the presence of the signature of open-ended evolution in this simulated world. The results show that the criteria for open-ended evolution are satisfied.

The rest of the paper is structured as follows. First we consider the modeling of open-ended evolution of natural systems. Next we describe the proposed cellular automata world. This is followed by the measuring and assessing the open-endedness of the proposed cellular automata world. Finally, the paper is closed by the discussion and conclusions section.

## Modeling open-ended evolution of natural systems

Natural systems continually take up and discard components from and into their environment. For example, consider molecules in the case of cells employees working for a company. The taken-up components undergo a series of interactions with other components and may become part of larger structural components of the system. For example, the molecules that are taken up may be broken into parts and some of these parts may be incorporated into other molecules. Or a new employee may undergo a series of trainings, get experience in a number of roles for some period, before joining one of teams and later may move on to a different role in a different team. The discarded components, e.g. waste molecules or employees who are fired, interact with further potential system components in the environment and may get taken up by another system. In general, all components of natural systems get replaced, for example on average a water molecule stays inside of the human body for about two weeks (Alberts et al, 2008) and all employees of a company get replaced or lose their role after some years, including the CEO. Thus these systems exist in the flux of components that temporarily make part of the system.

An individual example of a natural system is made of many components that interact and through patterns of such interactions deliver the behaviors of the individual. For example, proteins and other molecules interact in cells, cells interact in organisms, and organisms interact in social systems. The existence of an individual system can be seen as the maintenance of these interactions between components in accordance with individual specific rules about what interactions and patterns of interactions are useful and what are not useful for the individual. The components change, but the rules of interactions are maintained and followed through many component interactions throughout the life time of the individual system.

The individual systems do not take up randomly components from the environment, but select those that are required to maintain the functioning of the system. Through the interactions of components the individual system is able to make this choice sufficiently correctly on average in order to maintain its existence. For example, a bacterial cell may pick up a range of molecules from its environment, some of which are useful (e.g. sugar) and some of which are not useful (e.g. penicillin). Somehow the cell is able to select those molecular interactions that are useful for it (e.g. the use of sugar to generate energy storing molecules) and eliminate that ones that are harmful (e.g. by breaking up penicillin molecules before they can block the activity of their target proteins) (Blair et al 2015). Similarly, companies hire in general people

who they need to work for them, and who can deliver on average the job that they are hired for.

This implies that in order to simulate real-like open-ended evolution the simulation of the behavior of individuals in terms of interactions between components of the individuals is important. Furthermore, the existence of individuals should depend on the production of such behaviors, i.e. the existence of individuals should be checked by considering their behaviors and assessing whether these behaviors are consistent with the existence of the individual.

This reasoning also implies that biological and social systems somehow through their behaviors decide which elements of the environment are appropriate and which are not for the maintenance of their existence. For example, cells operate the right kind of channel molecules to pick up the right kind of molecules and ions from their environment. When the environment changes and one nutrient (e.g. an amino-acid) is replaced by another as abundant resource, the cells adapt and change their molecular pick-up behavior in order to use the available resource (Hottes et al 2013).

To adapt to environmental changes these systems must in some way predict their environment. For example, the cell that is ready to pick up one kind of nutrient from its environment experiences that the expected nutrient is not available and it also experiences other molecular interactions that indicate the presence of alternative nutrients. In response the cell changes the molecules exposed on its membrane that are in charge of facilitating the pick-up of nutrients. Such changes in the molecular composition of the cell membrane predict the cell's expectations about its environment.

Thus, in a sense, the natural systems compute their expectations about their environment. This computation happens through the interactions of their components. This means that in order to maintain their own existence these systems behave as autocatalytic systems that catalyze the reproduction (or recruitment) of correct components and interactions within themselves. They do this through a self-referential computation aimed to predict their environment (Andras, 2011).

The self-referential computation requires providing references to past interactions and components. This can be satisfied if all patterns of interactions (reference-able computations) can be represented by component (e.g. molecules that formed through corresponding molecular interactions), and if all patterns of components (reference-able data) can be represented by ongoing component interactions (e.g. molecular interactions which can happen only if the referenced pattern of molecules was present earlier). This circular referencing may appear irresolvable; however there is a mathematical formalism that can provide a solution, which is the theory of recursive domain equations (Pierce, 1991).

To put this more abstractly, systems that can produce open-ended evolution in nature must provide a practical realization of a solution of the following recursive domain equation:

$$R \cong A + [R \rightarrow R] \quad (1)$$

where  $R$  is a domain (e.g. a set or a category),  $A$  is a part of  $R$  and  $[R \rightarrow R]$  are all transformations of  $R$  to  $R$ , i.e. functions from  $R$  to  $R$ . In this formalism the patterns of components that are represented by  $A$  are not part of the system, but these exist only in the environment of the system.

The simplest non-empty solution of the recursive domain equation is the category of pre-orders, which is also a model of the  $\lambda$ -calculus (Pierce, 1991). This indicates that any natural system that is able to sustain open-ended evolution is such that it constitutes a representation of a solution of equation (1) and consequently it is also equivalent with a representation of the  $\lambda$ -calculus or equivalently of Turing machines. This means that these systems can (at least in principle) compute anything computable and predict their environment as much as it is predictable.

However, given the constraint of finite time available for computations about the environment these systems approximate in practice the prediction of their environment. Their ability to approximate their environment precisely depends on how efficient they are in terms of implementation of the solution of the recursive domain equation.

Thus in principle any system that aims to simulate the open-ended evolution that can be observed in nature must be able to produce a representation of a non-empty solution of equation (1). This means that these systems must be at least of the size of the category of pre-orders, which is comparably infinite as the category of sets, more infinite than the set of real numbers.

In order to deal with apparent difficulty of representing so infinite systems let us consider first real systems, such as cells or social systems. While the above argument implies that the cell must be an infinite system, the reality is that it is made of a finite set of molecules and molecular interactions at any time. To accommodate the infiniteness requirement, let us consider the life of a bacterial cell. The cell emerges after a division of another bacterial cell and it lasts until its own division into daughter cells. However, if we consider that the cell is in fact the continuation of the parent cell and its daughter cells are continuations of the cell itself, and take the whole life trajectory of the continuations of the cell, both backward and forward, we find that we are dealing with an infinite system. The number of kinds of molecules involved in cells is also similarly infinite, given the possible variations of molecules (e.g. consider the huge DNA molecules). So, while a given cell at any time provides a finite snapshot of the infinite cell system, considering the cell system in its totality, the system is indeed infinite (Andras, 2011).

This means that in order to be able to simulate a real-like system with open-ended evolution the system must be able to extend infinitely in principle and the simulation at any time should provide a finite snapshot of the infinite system, which represents a solution of the recursive domain equation. The key aspect is that the system must be extendable infinitely into a representation of the  $\lambda$ -calculus or equivalently of Turing machines. We note that somewhat similar ideas are presented in Hernandez-Orozco et al (2016), which also imply the infiniteness of systems able to produce open-ended evolution.

## A cellular automata world

Turning the reasoning of the previous section around, if we consider a system that explicitly or implicitly implements  $\lambda$ -calculus then we should be able to use this system to produce a simulation of open-ended evolution. Given that some cellular automata are equivalent of Turing machines (Wolfram, 2002), in principle these could be used to

implement such worlds of abstract systems, where a form of open-ended evolution may emerge and work. For this purpose we may consider Wolfram's rule 110 for one-dimensional cellular automata (Wolfram, 2002) in combination with random input strings of black and white squares. Here we describe a such one dimensional cellular automata world where open-ended evolution may emerge. We note that the cellular automata systems described in Adams et al (2017), in particular Case III, are similar to some extent to our proposed cellular automata worlds.

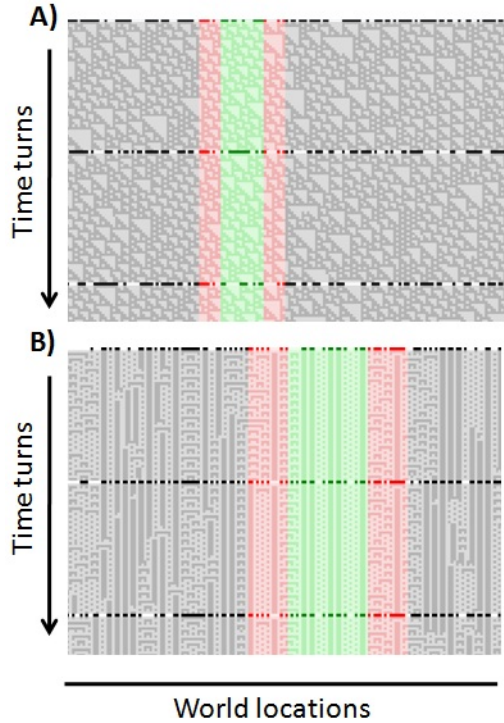
Our cellular automata world in principle is an infinite string of bits that can be either 1 or 0, the bits of the world are denoted as  $w(n), n \in \mathbb{Z}$  (in practice our simulated worlds are strings of 5 million bits). The world gets updated in each time turn according to the one dimensional cellular automata rule  $R(r)$ , where  $r \in \{0, \dots, 255\}$ . The rule is defined as follows: for a world bit  $w(n)$  the values of the bits to the left and right are  $w(n-1)$  and  $w(n+1)$ , the new value in the next time turn is  $r_u$  where  $u = w(n-1) + 2 \cdot w(n) + 4 \cdot w(n+1)$ ,  $u \in \{0, \dots, 7\}$  and the binary representation of  $r$  is  $r_0 r_1 \dots r_7$ . The world starts with randomly set bits, with equal probability for each bit to be 0 or 1. After each time turn a small proportion ( $\epsilon \ll 1$ ) of the world bits are randomly flipped.

The individuals in this world are bit strings. Each individual belongs to a species and all individuals belonging to a species have the same characteristic bit string. Each species  $s$  has a specific length  $l_s > 0$  and the characteristic bit string of the species is  $b_s = b_s(1), \dots, b_s(l_s)$ , such that  $l_s$  is always an even number. An individual  $d$  belonging to species  $s$  occupies  $2l_s$  bits in the world. Let the individual  $d$  have the starting (left-most) position in the world at  $d_{start} = n$ , then the world bits representing the individual are  $w(n), \dots, w(n+2l_s-1)$  and the values of these bits are set as follows:  $w(n+i-1) = b_s(i)$  for  $i = 1, \dots, l_s/2$ ,  $w(n+3l_s/2+i-1) = b_s(l_s/2+i)$  for  $i = 1, \dots, l_s/2$ . The values of the bits  $w(n+i)$  for  $i = l_s/2, \dots, 3l_s/2-1$  are not set specifically, but these are considered to be the inside of the individual. Overlap between individuals is not allowed.

A world turn is made of  $T^* (= 50)$  time turns. After each world turn the contents of the inside of individual is checked. If the bit series of the inside is sufficiently similar to the species specific bit string of the individual then the individual survives – the ratio of matching bits relative to the length of the species bit string is  $z$  and the survival criterion is  $z > z_{survive}$ . The  $z_{survive}$  value may change depending on how over-populated or under-populated is the world (the initial  $z_{survive} = 0.37$ ). If an individual survives after a world turn the set bits of the individual are set again to the species specific values, while the bits in the inside of the individual are left unchanged. If an individual does not survive it dies and the bits allocated to this individual become part of the environment and the species specific part of the individual is not set to the specific bit values.

The  $z_{survive}$  value changes adaptively to prevent excessive population growth and also excessive population shrinking. In our implementation of the cellular automata world  $z_{survive}$  increases if the total population is above an upper limit and the increases get larger if the population continues to grow. Similarly, if the population size is below a lower limit the value of  $z_{survive}$  decreases and the decrease gets quicker if the population continues to shrink. A practical impact of

this in our implementation of the cellular automata world is that the size of the population of the individuals often oscillates between values close to the lower and upper limits which trigger the changes in the  $z_{survive}$ .



**Figure 1.** Two examples of individuals in cellular automata worlds – A) Rule 110; B) Rule 73. The rows with the darker colors are the rows corresponding to world turns. The red and green colored parts are the species specific and inside parts of the individuals, the grey, black and white colored part is the environment. In A) the environment has much more impact on the inside of the individual than in B).

During the world turn the rule of the cellular automata world is applied to calculate the new bit values for the world. Through this process the environment outside of the individual influences the changes that happen in the inside of the individual. Similarly, the individual also influences what happens outside of itself, in its world environment. In this way the individuals in this world get influenced by their environment and influence their environment.

If the ratio  $z$  of the matching bits in the inside of the individual is sufficiently high,  $> z_{reproduce} = 0.63$ , the individual produces an offspring. The offspring production may double the individual's bit string, by concatenating two copies of the species specific bit string (with probability  $p_{double} = 0.005$ ). The offspring may inherit a split bit string, by keeping the first half of the species specific bit string and possibly one additional bit, if the resulting bit string length would be odd (with probability  $p_{split} = 0.005$ ). It is also possible that two reproduction ready individuals produce a joint offspring by concatenating their bit strings (with probability  $p_{join} = 0.01$ ). In most of the cases the offspring simply inherits the bit string of the parent individual. To generate the bit string for the offspring the inside bit string of the individual(s) is used with corrections, i.e. the non-

matching bits are corrected to match the correct bit value with  $z_{correction} = 0.931$  probability. The new individual is placed randomly into a part of the world that does not overlap with the location of any other existing individual. The generated new individual is checked to determine its species identity. If the species specific bit string of the individual matches closely the bit string of an existing species ( $z_{same} = 0.9$ ), the individual is assigned to that species. If this is not the case a new species emerges in the world and the specific bit string of this new species is set to match exactly the species specific bit string of the new individual.

Through the evolution of the world new species are created, species may die out and the number of individuals belonging to a species varies. The dominant species and the number of individuals belonging to this species also change. The dynamics of the world may change significantly if the rule that defines the world is changed. The reported simulations were run mostly using rule  $R(110)$  and  $R(73)$ . Figure 1 shows an illustrative example of the evolution of the proposed cellular automata world for two worlds corresponding to two rules ( $R(110)$  and  $R(73)$ ).

### Measuring open-ended evolution in the cellular automata worlds

To measure the open-endedness of evolution of species in the cellular automata world we follow the approach proposed by Bedau (1992; Bedau et al, 1998). A shadow evolutionary system is set up in parallel with the evolution in the cellular automata world. The shadow world is represented as a list of shadow individuals belonging to shadow species. Every time an individual dies in the cellular automata world a randomly picked shadow individual dies. Every time a new individual is created a shadow individual is created belonging to a randomly picked shadow species. When a new species emerges in the cellular automata world a new shadow species is created and the last newly created individual is set to belong to this new shadow species. The shadow is used to calculate the same evolutionary metrics that are calculated for the cellular automata world and the metrics are compared to assess the relevance of the evolutionary process. In general it is expected that the world with the evolutionary process differs significantly from the shadow world in terms of these metrics and shows stronger evidence indicating the open-endedness of the world than the shadow world.

The open-endedness metrics that we use are based on the metrics of Bedau (1992; Bedau et al, 1998) and are as follows. Diversity at world turn  $t$  is the number of species with active individuals

$$D(t) = \#\{s_i | \exists v_k(t) \in s_i; i = 1, \dots, N_s(t); k = 1, \dots, N_v(t)\} \quad (2)$$

where  $s_i$  are species identifiers,  $v_k(t)$  are identifiers of individual active at world turn  $t$ ,  $N_s(t)$  is the total number of species that have existed up to world turn  $t$  in the world,  $N_v(t)$  is the number of live individuals in the world at world turn  $t$ ,  $v_k(t) \in s_i$  indicates that the individual  $v_k(t)$  belongs to species  $s_i$  and  $\#$  denotes the cardinality of the set. For the shadow world diversity is defined similarly, as

$$D'(t) = \#\{s'_i | \exists v'_k(t) \in s'_i; i = 1, \dots, N_{s'}(t); k = 1, \dots, N_{v'}(t)\} \quad (3)$$

where the prime notation indicates the shadow world.

The components in the cellular automata world are the species. The weighted component activity of the world is calculated by considering all individuals belonging to each species with live representatives. The component activity for each species is the amount of time measured in world turns for which the species has been active in the sense of having live individuals representing the species. We use the weighted world turn age of the species, where the weight is the number of live individuals belonging to the species. The reason behind this approach is that each current individual instance of a species represents a line of existence of the species since the origin of the species. The average weighted component activity of the world is the ratio between the sum of weighted component activities and the number of active species in the world (i.e. the diversity of the world). The average weighted component activity of the world is defined as follows:

$$\alpha_i(t) = \begin{cases} 1 & \text{if } \exists v_k(t) \in s_i; k = 1, \dots, N_v(t) \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

$$a_i(t) = \begin{cases} \sum_{\tau=1}^t \alpha_i(\tau) & \text{if } \alpha_i(t) > 0 \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

$$m_i(t) = \#\{v_k(t) | v_k(t) \in s_i; k = 1, \dots, N_v(t)\} \quad (6)$$

$$c_i(t) = a_i(t) \cdot m_i(t) \quad (7)$$

$$\bar{C}(t) = \frac{\sum_{i=1}^{N_s(t)} c_i(t)}{D(t)} \quad (8)$$

where  $\alpha_i(t)$  is the existence indicator of the species  $s_i$  at world turn  $t$ ;  $a_i(t)$  is the length of active existence (or age) in terms of world turns for the species  $s_i$  at world turn  $t$ ;  $c_i(t)$  is the total component activity for species  $s_i$  at world turn  $t$ ;  $\bar{C}(t)$  is the average weighted component activity of the world at world turn  $t$ ; the other notations are the same as for the previous equations. The corresponding equations, definitions and notations work for the shadow world as well, in a similar manner as equation (3) for the shadow world corresponds to equation (2) for the cellular automata world – the average weighted component activity of the shadow world is denoted as  $\bar{C}'(t)$ .

The innovation activity in the cellular automata world is represented by those species that have existed for a sufficient long time, but still they are sufficiently recent as well. A new species may emerge, but it may also die out in a few world turns, indicating that it does not represent an evolutionarily advantageous innovation. If the species is present already for a long time in the world then it no longer represents an evolutionary innovation. For practical purposes we set the existence interval that is considered for innovation activity from 10 world turns to 20 world turns. Similarly as above, the age of such species is multiplied with the number of individuals belonging to the species to get the total weighted innovation activity by the species. The average weighted innovation activity in the cellular automata world is given by the ratio of the sum of total weighted innovation of species and the number of active species (diversity). The equations

defining the average weighted innovation activity of the world are as follows.

$$e_i(t) = \begin{cases} a_i(t) \cdot m_i(t) & \text{if } T_0 \leq a_i(t) \leq T_1 \\ 0 & \text{otherwise} \end{cases} \quad (9)$$

$$\bar{E}(t) = \frac{\sum_{i=1}^{N_s(t)} e_i(t)}{D(t)} \quad (10)$$

where  $T_0$  and  $T_1$  are the lower and upper limits in terms of world turns for a species to be considered an innovation in the world ( $T_0 = 10, T_1 = 20$ ),  $e_i(t)$  is the total weighted innovation for species  $s_i$  at world turn  $t$ ;  $\bar{E}(t)$  is the average weighted innovation of the world at world turn  $t$ ; the other notations are the same as for the previous equations. The average weighted innovation for the shadow world is defined in a similar manner – this is denoted as  $\bar{E}'(t)$ .

In addition to measuring the diversity, average weighted component activity and average weighted innovation in the world we also used two other measures to assess the difference between the cellular automata world and its shadow. These measures are the frequency of change of the dominant species and the ratio between the number of individuals belonging to the dominant species and the total number of individuals in the world. The additional measures are defined in terms of equations as follows.

$$i^*(t) = \operatorname{argmax}_i m_i(t) \quad (11)$$

$$h(t) = \begin{cases} 1 & \text{if } i^*(t) \neq i^*(t-1) \\ 0 & \text{otherwise} \end{cases} \quad (12)$$

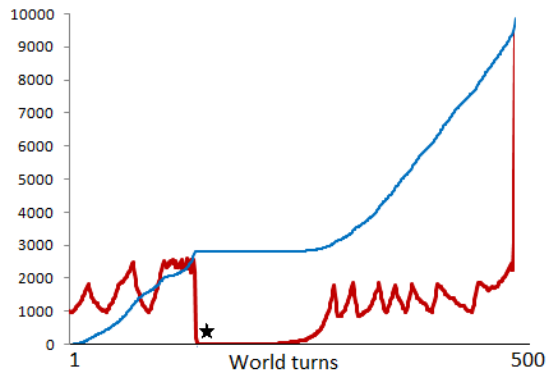
$$\varphi(t) = \frac{\sum_{\tau=1}^t h(\tau)}{t} \quad (13)$$

$$\rho(t) = \frac{m_{i^*(t)}}{N_v(t)} \quad (14)$$

where  $i^*(t)$  is the index of the species with the maximal number of individuals at world turn  $t$ , i.e. the dominant species;  $\varphi(t)$  is the frequency of changes of dominant species at world turn  $t$ ; and  $\rho(t)$  is the ratio between the number of individuals belonging to the dominant species and the total population of the world. The matching definitions apply to the shadow world and the shadow frequency of changes is denoted as  $\varphi'(t)$  and the population ratio for the dominant shadow species is  $\rho'(t)$ .

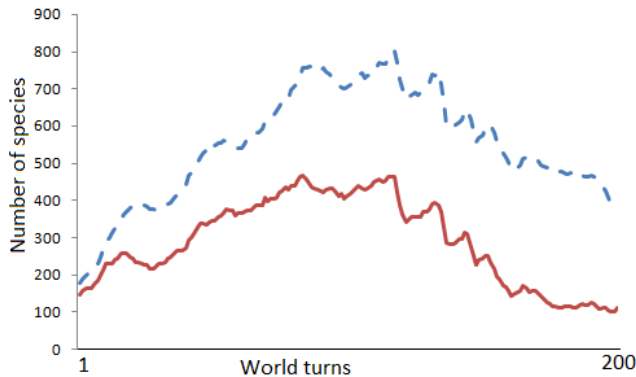
The simulated worlds that we generated in many cases had drastic collapses of the number of individuals, following the steady and often rapid increase in the size of the total population of individuals. Such events can be seen as mass extinction events in the simulated world, which are followed by the regeneration of the diversity of the species (in most cases) or alternatively the complete die-out of all species (in some of the cases). Figure 2 provides an example of the evolution of the population size in one simulation which experienced mass extinction and rebounding of the species diversity. The mechanistic cause of these events in our simulated worlds is the increase of the  $z_{survive}$  value, which is triggered by the large and increasing size of the  $T_0$  to appreciate the volume of evolutionary activity, Figure 2 also shows the cumulative number of new species.

We found that in all simulations the diversity of the shadow world was larger in general than the diversity of the cellular automata world (see Figure 3 for an example). The reason of



**Figure 2.** Typical evolution of the population size (thick red line) and of the cumulative number of new species (thin blue line) in the cellular automata world, including oscillating population size and a mass extinction event marked by the star.

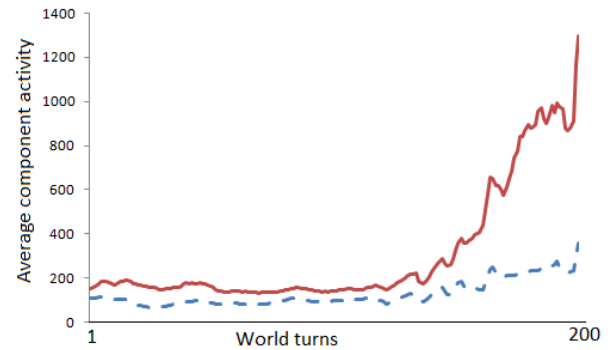
this is that every time when a new species is created in the cellular automata world a new species is also created in the shadow world, however the species die out through evolutionary selection in the cellular automata world and randomly in the shadow world. Thus, while the least fit species are likely to die out quickly in the cellular automata world the corresponding species may last considerably longer in the shadow world.



**Figure 3.** Typical evolution of the number of live species (diversity) in the cellular automata world (continuous red line –  $D(t)$ ) and the shadow world (segmented blue line –  $D'(t)$ ).

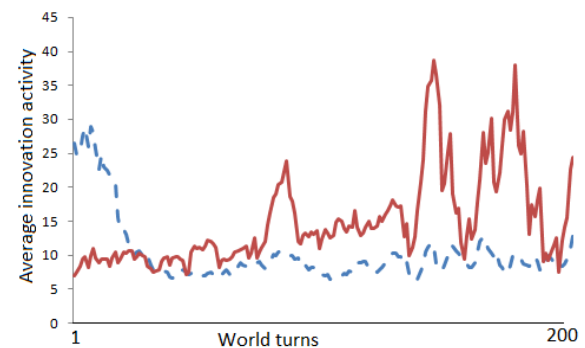
The average component activity in most cases is higher in the case of the cellular automata world than in the case of the shadow world – see Figure 4. This implies that successful species persist for longer in the cellular automata world than in the shadow world and they also have more copies in the proper world than in the shadow world. This is due to the evolutionary selection of the best performing individuals that belong most likely to the most successful species. This counter balances the higher diversity in the shadow world and shows that the evolution in the cellular automata world is driven by the evolutionary process and not some form of random selection. The average component activity of the cellular automata world grows in general and this growth stops only in cases of larger or smaller scale extinction events (i.e. when the number of individuals and species drastically

drops, possibly to the level of less than a handful species with only a few representative individuals for each – see Figure 2).



**Figure 4.** Typical evolution of the average weighted component activity in the cellular automata world (continuous line –  $\bar{C}(t)$ ) and the shadow world (segmented line –  $\bar{C}'(t)$ )

The average innovation activity in the cellular automata world is usually higher than the innovation activity in the shadow world, although not always (see Figure 5). Typically the average innovation activity follows a growth trend, although the actual variation in the activity has both ups and downs. This shows that the cellular automata worlds proposed here generate sufficiently persistent novel species following an increasing trend and in general much more than what is generated randomly in the shadow world. Naturally, the amount of innovation gets reduced following major extinction events. The drivers of the innovation are the reproduction mechanisms described above, i.e. imperfect copying of the parent's species specific bit string, doubling, splitting and joining of bit strings, which lead to the emergence of new species in the cellular automata world.

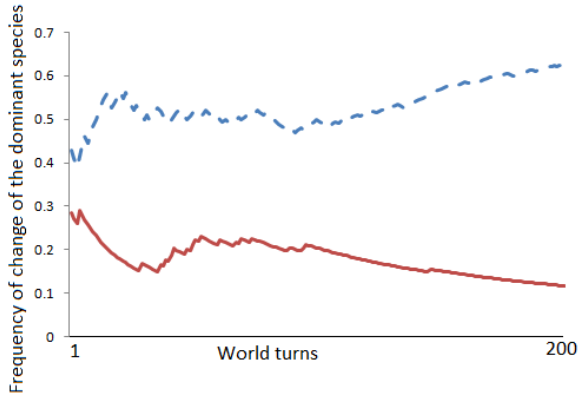


**Figure 5.** Typical evolution of the average weighted innovation activity in the cellular automata world (continuous red line –  $\bar{E}(t)$ ) and the shadow world (segmented blue line –  $\bar{E}'(t)$ )

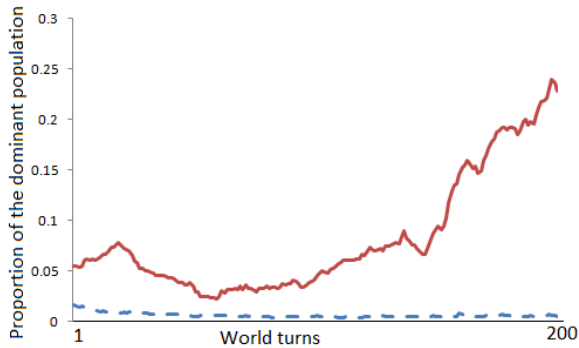
Comparing the frequency of changes of dominant species between the proper world and the shadow world shows that the dominant species changes much more frequently in the shadow world. This indicates that the dominant species in the cellular automata world are determined by evolutionary selection and not randomly as in the shadow world. Similarly, the comparison of the proportions of the dominant species between the proper and shadow worlds show that the dominant species are much more dominant in the proper



world. Again this confirms the importance of the evolutionary selection in the cellular automata world in comparison with the random determination of the dominant species in the shadow world. Figures 6 and 7 show an example for the comparison of frequencies of dominant species changes and proportions of dominant species for the cellular automata world and the shadow world.



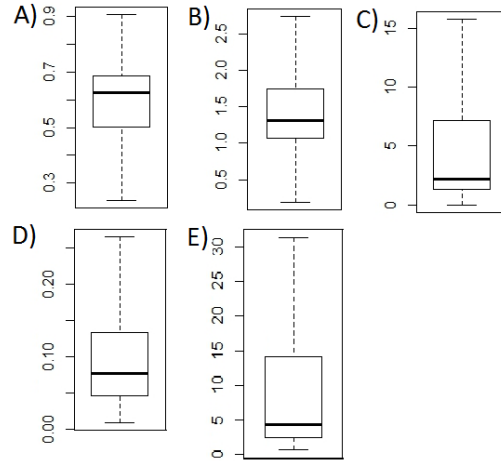
**Figure 6.** Typical evolution of the frequency of change of the dominant species in the cellular automata world (continuous red line –  $\varphi(t)$ ) and the shadow world (segmented blue line –  $\varphi'(t)$ ).



**Figure 7.** Typical evolution of the proportion of the dominant species in the cellular automata world (continuous red line –  $\rho(t)$ ) and the shadow world (segmented blue line –  $\rho'(t)$ ).

The distributions of the values of the ratios between species numbers (diversity), average weighted component activities, average weighted innovation activities, frequencies of change of dominant species and proportions of dominant species for the cellular automata world and the shadow world are shown in box plot format in Figure 8. The data for these distributions represents five different runs of the simulations containing over 1,200 data values. This analysis confirms that the diversity of the cellular automata world is smaller than that of the shadow world – the typical ratio range is 0.5 – 0.69 for  $D(t)/D'(t)$ . It shows that average weighted component activity in the cellular automata world is above this activity of the shadow world with the typical ratio of  $\bar{C}(t)/\bar{C}'(t)$  being above 1.07 and below 1.74. The weighted innovation activity in the cellular automata world can be much larger than the same activity in the shadow world and the typical values for

the ratio  $\bar{E}(t)/\bar{E}'(t)$  are in the range of 1.35 – 7.21. The dominant species changes much less frequently in the cellular automata world than in the shadow world, with typical ratios for  $\varphi(t)/\varphi'(t)$  in the range of 0.04 – 0.14. On the other side, the proportion of the dominant species in the total population is much larger in the cellular automata world than in the shadow world with typical ratios for  $\rho(t)/\rho'(t)$  being in the range of 2.57 – 14.11.



**Figure 8.** Box plots of the value distributions for the ratios between: A) the species numbers –  $D(t)/D'(t)$ ; B) the average component activities –  $\bar{C}(t)/\bar{C}'(t)$ ; C) the average innovation activities –  $\bar{E}(t)/\bar{E}'(t)$ ; D) the frequencies of change of the dominant species –  $\varphi(t)/\varphi'(t)$ ; E) the proportions of the dominant species –  $\rho(t)/\rho'(t)$ . Outliers are not shown to avoid clutter in the figure.

The results reported here show that the evolutionary selection process plays a defining role in the evolution in the cellular automata world. Our data shows that the evolution of the average weighted component activity and weighted innovation activity follow a growth path in general. The growth path of average component activity is broken occasionally by extinction events imposed by the growth constraints that we implemented in the cellular automata world. The average innovation activity has large variations along the growth path. The observed growth paths of the average weighted component activity and innovation activity indicate that the proposed cellular automata world has the potential for open-ended evolution.

## Discussion and conclusions

The simulations reported here typically stop after 200 – 800 world turns due to saturation of the simulated world with individuals. To appreciate the volume of evolutionary activity, consider that in the example shown in Figure 2 the average number of new species per world turn is 200 and in a slowed down version of cellular automata world simulations the average number of new species per world turn is around 1. This implies that each world turn in the reported simulations is equivalent of 10,000 (50 x 200) time turns in the slower pace simulation. So, 200 – 800 world turns in the reported simulations are in effect equivalent of 2 – 8 million time turns in slower simulations of the cellular automata world.

We chose in this paper to report the weighted versions of the component and innovation activity. The fundamental reason for this is explained above. One may object that the increase in the weighted version may be due purely to the increase in  $m_i(t)$ , the count of individuals belonging to possibly very short lived species. However, this value can only increase as new individuals are created over time, and the number of newly created individuals belonging to a species is at most the number of existing individuals (usually below this number as not all individuals reproduce). Typically the number of individuals belonging to a species first grows and then falls as fitter species emerge. The weighted measures capture better the evolutionary superiority of the cellular automata world compared to their shadow worlds. Due to the creation of a shadow species every time when a real species is created, and the random selection of shadow species for adding and eliminating individuals, the average age of shadow species is larger than the average age of real species. The individuals of the latter are selected for reproduction and death on the basis of their performance in an evolutionary competition, which leads to the elimination of the less well performing species relatively quickly. Thus, while the average age of species increases in general for both the real and shadow world, it is higher for the shadow world. This may be read misleadingly as an indication that the shadow world has more evolutionary activity, which is wrong, as shown by the measures reported in the paper.

To summarize, the proposed cellular automata worlds show positive average component activity and average innovation activity along the evolutionary trajectory of the world and both measures of open-endedness follow an increasing trend over time. This indicates that these cellular automata worlds are able to produce open-ended evolution according to these measures (Bedau, 1992; Bedau et al, 1998). The comparisons with the associated random shadow worlds show that the evolutionary selection processes play a defining role in the cellular automata worlds and that their evolution is not similar to a random selection process.

In our introductory reasoning we noted the importance of the Turing equivalence of the underlying mechanism of the cellular automata world. We chose to use a world rule (110) that is assumed to satisfy this requirement. However, we have not assessed the nature and the extent to which this feature of the world influences the evolutionary process in the world.

We believe that the proposed cellular automata worlds provides a novel avenue for the research on open-ended evolution and the potential and features of evolutionary processes. While the proposed worlds share similarity with other cellular automata based approaches (e.g. Sayama, 1998, 1999; Adams et al, 2017), they are simpler (one dimensional), infinite and mechanistically more transparent than higher dimensional cellular automata based world.

The source code (in Delphi) for the implementation and simulation of the proposed cellular automata worlds is available on request from the author.

## References

Adams, A., Zenil, H., Davies, P.C.W., Walker, S.I. (2017). Formal definitions of unbounded evolution and innovation reveal universal

- mechanisms for open-ended evolution in dynamical systems. *Scientific Reports*, 7:997.
- Alberts, B. et al (eds.) (2008). *Molecular Biology of the Cell*. Garland Science, 5<sup>th</sup>ed.
- Andras, P. (2011). Modelling living systems. Kampis, G., Karsai, I, Szathmary, E., editors, *ECAL 2009*, pages 706-713, Springer, Heidelberg.
- Bedau, M. (1992). Measurement of evolutionary activity. Langton, C.G. Taylor, C., Farmer, J.D., Rasmussen, S.E. editors, *Artificial Life II*, pages 431-461, Addison-Wesley, Boston, MA.
- Bedau, M.A., Snyder, E., Packard, N.H. (1998). A classification of long-term evolutionary dynamics. Adami, C., Belew, R.K., Kitano, H., Taylor, C.E., editors, *Artificial Life VI*, pages 189-198, MIT Press, Cambridge, MA.
- Blair, J.M.A. et al. (2015). Molecular mechanisms of antibiotic resistance, *Nature Reviews*, 13:42-51.
- Channon, AD (2006). Unbounded evolutionary dynamics in a system of agents that actively process and transform their environment. *Genetic Programming and Evolvable Machines*, 7:253-281.
- Channon, AD, Damper, RI (1998). Evolving novel behaviours via natural selection. Adami, C., Belew, R.K., Kitano, H., Taylor, C.E., editors, *Artificial Life VI*, pages 384-388, MIT Press, Cambridge, MA.
- Fortuna, MA, Zaman, L, Ofria, C, Wagner, A (2017). The genotype-phenotype map of an evolving digital organism. *PLoS Computational Biology*, doi: 10.1371/journal.pcbi.1005414.
- Hernandez-Orozco, S, Hernandez-Quiroz, F, Zenil, H (2016). Undecidability and irreducibility conditions for open-ended evolution and emergence, *ArXiv*, arxiv.org/abs/1606.01810.
- Hottes, A.K. et al (2013). Bacterial Adaptation through Loss of Function. *PLoS Genetics*, 9(7): e1003617.
- Knibbe, C, Mazet, O, Chaudier, F, Fayard, J-M, Beslon, G (2007). Evolutionary coupling between the deleteriousness of gene mutations and the amount of non-coding sequences. *Journal of Theoretical Biology*, 244:621-630.
- Langton, CG (1984). Self-reproduction in cellular automata. *Physica D*, 10:135-144.
- Langton, CG (ed.) (1995). *Artificial Life. An Overview*. MIT Press, Cambridge, MA.
- Luhmann, N (1996). *Social Systems*. Stanford University Press, Stanford, CA.
- Ofria, C, Wilke, CO (2006). Avida: A software platform for research in computational evolutionary biology. *Artificial Life*, 10:191-229.
- Pierce, B.J. (1991). *Basic Category Theory for Computer Scientists*. MIT Press, Cambridge, MA.
- Ray, TS (1991). An approach to the synthesis of life. Langton, C.G. Taylor, C., Farmer, J.D., Rasmussen, S.E. editors, *Artificial Life II*, pages 371-408, Addison-Wesley.
- Sayama, H (1998). Introduction of structural dissolution into Langton's self-reproducing loop. Adami, C., Belew, R.K., Kitano, H., Taylor, C.E., editors, *Artificial Life VI*, pages 114-122, MIT Press, Cambridge, MA.
- Sayama, H (1999). A new structurally dissolvable self-reproducing loop evolving in a simple cellular automata space. *Artificial Life*, 5:343-365.
- Shao, J, Ray, TS (2010). Maintenance of species diversity by predation in the Tierra system. Fellerman, H., et al., editors, *Artificial Life XII*, pages 533-540, MIT Press, Cambridge, MA.
- Soros, LB, Stanley, KO (2014). Identifying necessary conditions for open-ended evolution through the artificial life world of Chromaria. Sayama, H., Rieffel, J., Risi, S. Doursat, R., Lipson, H., editors, *Artificial Life 14*, pages 793-800, MIT Press, Cambridge, MA.
- Wolfram, S. (2002). *A New Kind of Science*. Wolfram Media, Champaign, IL.
- Yaeger, LS (1994). Computational genetics, physiology, metabolism, neural systems, learning, vision, and behavior. Langton, C.G., editor, *Artificial Life III*, pages 263-298, Addison-Wesley, Boston, MA.
- Yaeger, LS, Sporns, O (2006). Evolution of neural structure and complexity in a computational ecology. Rocha, L.M. et al, editors, *Artificial Life X*, pages 330-336, MIT Press, Cambridge, MA.