

PAPER • OPEN ACCESS

A heuristic neural network model in the research of properties of evolutionary trajectories

To cite this article: S Bartsev and P Baturina 2019 *IOP Conf. Ser.: Mater. Sci. Eng.* **537** 042001

View the [article online](#) for updates and enhancements.



IOP | ebooks™

Bringing you innovative digital publishing with leading voices to create your essential collection of books in STEM research.

Start exploring the collection - download the first chapter of every title for free.

A heuristic neural network model in the research of properties of evolutionary trajectories

S Bartsev^{1,2} and P Baturina²

¹Institute of Biophysics SB RAS, Federal Research Center, Krasnoyarsk Scientific Center SB RAS, 50, Akademgorodok, Krasnoyarsk, 660036, Russian Federation

²Siberian Federal University, 79 Svobodny pr., Krasnoyarsk, 660041, Russian Federation

E-mail: bartsev@yandex.ru

Abstract. There is considerable data on molecular evolution, but there remains no approach to systematizing them within the framework of the key problems of biology. To search for the most common properties of evolving systems, the heuristic method has been proposed. Artificial networks of formal neurons were chosen as the heuristic model object. The paper examines the divergent component of evolutionary trajectory formation. As a result of the simulation, the dependence of the potential variability parameter on the position of the fitness function landscape was obtained. The simulation results are in agreement with the real data of molecular evolution experiments.

1. Introduction

The question of whether evolution proceeds along a deterministic path or can proceed along alternative trajectories has attracted many researchers [1], [2], [3], [4], [5], [6]. Experimental evolutionary studies in conjunction with the mathematical modelling of fitness function landscapes provide the opportunity to work with the problem of reproducibility of evolutionary trajectories.

However, the success of research depends largely on how adequate the concept of the process under study is. Currently, there are some contradictions between certain theoretical concepts and experimental data. For example, in the paper with the expressive title 'Replaying the tape of life: quantification of the predictability of evolution' [7], the only possibility is considered, namely that evolution trajectories converge to the one fittest structure of the evolving biological system. In this way, if the mutation rate is low then the evolution of a population goes along a single trajectory on the fitness landscape. In the case of a high mutation rate, the population explores multiple trajectories in parallel. In both cases, the final point of the evolution is assumed to be single.

At the same time, there are experimental data demonstrating the existence of several final points of evolution [4], [5], [8], [9]. It means that the concept of one fittest structure must be revised. The difficulty of studying in this field is associated with the deficiency of experimental data on evolution. Even an investigation of local evolution processes [4], [5] requires considerable effort and time.

There is a possibility to avoid this difficulty by referring to heuristic models. The construction of abstract models in order to discover useful concepts, broad principles and general theories was called the heuristic method by J. von Neumann [10]. The necessity to use artificial model objects when complex a biological system is being explored was stated by von Neumann: "Appealing to the organic,



living world does not help us greatly, because we do not understand enough about how natural organisms function. We will stick to automata which we know completely because we made them” [10].

The abstract heuristic model object is “transparent”, everything is known about its structure, and this allows efforts to be concentrated on identifying general properties, and not on finding seemingly important, but in fact infinite, details of the structure of real biological objects.

To choose the type of heuristic model which is able to represent the key properties of an evolving biological system, we have to identify these properties. We refer to J. Bernal who outlined the specificity of biological evolving systems [11]. He said that “biology is methodologically different from other natural sciences in that the focus is primarily on the functioning and evolution of systems. The structure here has meaning only in relation to function and origin. ... The origin, structure and function can no longer be separated”.

The heuristic model should have a simply described structure that implements a certain function that arose during the evolutionary process. Artificial neural networks (ANN) are well suited for playing the role of the heuristic model since it is easy for them to determine the function, select the structure and start the learning process, the formal description of which, according to key features, does not differ from the evolutionary process in biology.

Why can we consider the learning process of ANN similar to the evolution process of living beings? Evolution is the process of changing the structure of a system in accordance with a certain functional criterion of optimality [12], [13], [14], [15]. In a general case, the evolution process can be considered as a search for an extremum of some target function $H(a_j)$ in the presence of constraints corresponding to structural limitations of any kind. The problem of finding an extremum in the presence of constraints can be solved using the method of indefinite Lagrange multipliers through the introduction of the generating function:

$$W = H(\{\alpha_j\}, A_k) + \sum_i \mu_i \Phi_i(\{\alpha_j\}, A_k, \{\beta_m\}) \quad (1)$$

where α_j are some ‘phenotype’ variables; A_k are external influences; β_m are structural parameters influencing the ‘phenotype’; μ_i are indefinite Lagrange multipliers; Φ_i are functions describing structural constraints.

The argument in favour of the phenomenological equivalence of systems described by the formula (1) is the No Free Lunch (NFL) theorem [16]. The theorem says that the effectiveness of any two algorithms in the search for an extremum is on average (for all possible target functions) identical. From this theorem, it follows that the existence of living beings is in no sense an indicator of the efficiency of natural selection. That, in turn, means that any system that changes its structure in accordance with a certain principle of optimality can serve as a model of biological evolution.

The ease of obtaining ensembles of ANNs allows a multiplicity of evolutionary outcomes to be simulated, which makes it possible to apply statistical methods and comparative analysis to eliciting common properties of the fitness landscape of ANNs.

Despite the fact that ANNs are highly abstract models of a biological system, they enable us to achieve concreteness in describing properties that are difficult to formalize, such as ‘structure’, ‘function’ and ‘complexity’, which is very difficult to do in relation to living systems.

By means of heuristic neural network models, it has previously been shown that the ability of biological systems to evolutionary changes is provided by the redundancy of the structure over the complexity of the function being formed, and as a result, with the fundamental multiplicity of evolutionary outcomes [17], [18], [19].

On the basis of these results, which are supported by recent experimental data [4], [5], we can formulate an evolutionary concept: the observed evolutionary changes are a combination of structural changes towards increasing the fitness function and changes that keep the functioning unchanged (conserving the functional invariant of the structure).

Thus, we follow the aim of the article [7] to find a quantitative measure of predictability as applied to the statistical ensemble of trajectories, but assuming the existence of a set of final endpoints instead of the assumptions about only one final endpoint.

2. Method and materials

Simple 3-neuron recurrent fully-connected ANNs were used in the investigation (figure1). The ANN functioning is described by

$$\alpha_i^{n+1} = \frac{\rho_i^n}{a + |\rho_i^n|}, \quad \rho_i^n = \sum_j x_{ij} \alpha_j^n + A_i^n, \quad (2)$$

where x_{ij} is the weight matrix; A_i^n is the input signal of the i -th neuron at the moment n ; α_i^n is the output of the i -th neuron at the n -th time; a is the coefficient defining the slope of the transfer function.

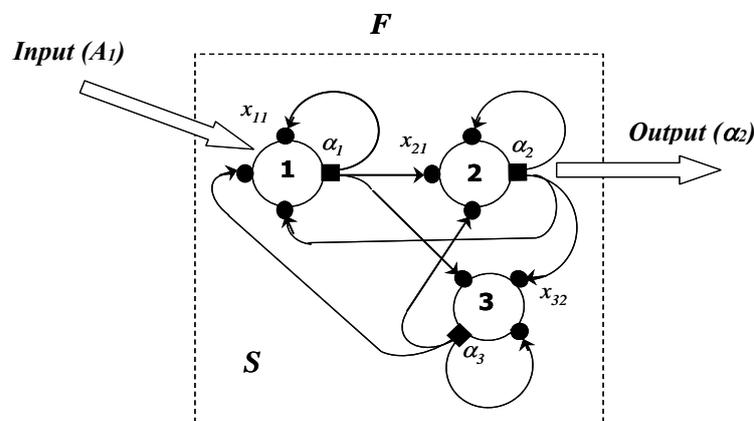


Figure 1. The scheme of the used neural network.

The box F describes functional mapping $F:A \rightarrow \alpha$, and S is the structure providing this mapping. As the measure of fitness the next target function was used:

$$H(\alpha_{out}^n, \delta_{out}^n) = \frac{1}{2} \sum_{out} (\alpha_{out}^n - \delta_{out}^n)^2 \quad (3)$$

where α_{out}^n is the signal on the output neurons at time n , δ_{out}^n is the signal required from the network at time n .

The formulae describing the functioning of the ANN play the role of structural constraints from Eq.1: $\Phi_i(\alpha_i^{n+1}, \{\alpha_j^n\}, A_i^n, \{x_{ij}\}) = \alpha_i^{n+1} - \rho_i^n / (a + |\rho_i^n|)$. As an illustration, the strict relationship between the generating function (Eq.1) and the learning Back Propagation algorithm was shown explicitly in [20]. For the sequential random search algorithm (imitating a mutation process) used in the study the feasibility of Eq.1 seems obvious. At each step of the training, a small random increment of weights was made and the value of the target function (3) was calculated. If the new target function value was less than previous one, then the new values of the weights were accepted and the process was repeated.

Three abstract simple functions were used for the training – a network has to respond to a fixed single impulse at the zero-time moment by some time pattern shown on the bottom row of figure2. In figure2 3D projections of 9D space of weight coefficients of ANN (upper row) for corresponding functions (bottom row) are shown. For clarity, it is necessary to emphasize that each coloured dot on the figures in the upper row represents the structure of one ANN performing the same function as the

ANNs represented by other dots of the same colour, and with the same fitness. It is a visual manifestation of the functional symmetry [17], [18], [19].

3. Results

The conducted experiments showed that trajectories of network learning in the structural space are prone to divergence. In figure.3A, some examples of the bundles of trajectories starting from one initial point and finishing at several endpoints are shown. Figure 3B demonstrates the divergence of trajectories from different points on the first ('mother') trajectory. It is clear that similar bundles of trajectories will arise from any point of this and other trajectories.

It is necessary to state that the divergence of the trajectories is a manifestation of the functional invariance of the neural network structure [17], [18], [19], which means that the observed evolutionary changes are a combination of changes in the structure towards the growth of the function of adaptation and a functional-invariant variation of the parameters, corresponding to neutral Kimura mutations [21].

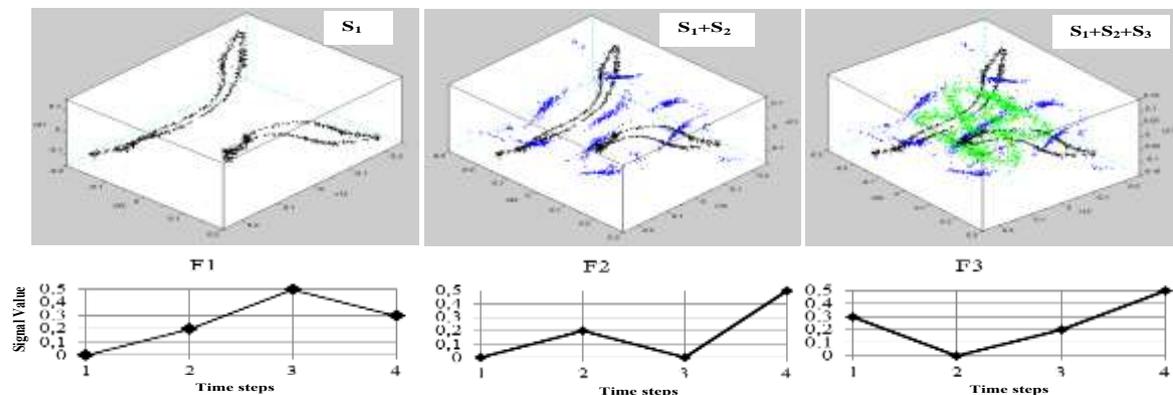


Figure 2. Configurations of equifunctional endpoint sets of neural network structures (upper row) corresponding to realized functions (bottom row).

Cumulative 3D projections of 9D weight space of neural network structures (S_i) that perform the same function (F_i) are shown. It is clearly seen that the structures of networks that perform the same function can differ more significantly than the structures of networks that perform different functions.

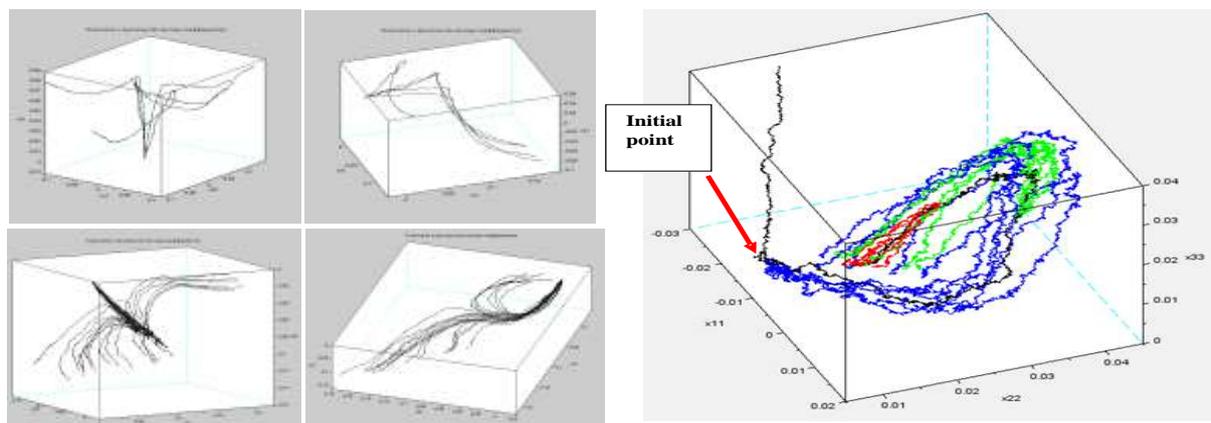


Figure 3. Some examples of trajectory divergence in the case of starting from one initial point (A) and the divergence of trajectories from a different point on the 'mother' trajectory (B).

Upper and bottom rows of A represent two different realizations, and the left and right columns correspond to different rotation views. Black on the picture (B) denotes two 'mother' trajectories that

emerged from the same point in the structural space of the neural network. Trajectories of a different colour are obtained by restarting training from three arbitrary points along one of the ‘mother’ trajectories.

A natural question arises regarding the properties of the landscape of the fitness function. In particular, does the potential variation of trajectories associated with a divergent contribution change in the course of an evolutionary movement towards maximum fitness? To answer this question, a series of experiments on the simulation of the divergence of trajectories was performed.

A quantitative estimation of the divergence from a point on the learning trajectory was introduced in the course of the work. The procedure of the estimation consists of the following steps (figure4):

1. From an estimated point on the ‘mother’ trajectory several (N) realizations of ANN training are made and the end positions of trajectory points after fixed numbers of training steps are stored.
2. A nine-dimensional vector (the coordinates of which are an approximation of the plot of the curve by the method of least squares) is assigned to each child trajectory.
3. The lengths of the obtained vectors are equalized since we are only interested in their angular divergence.
4. Then using the least squares method an average vector of the same length as the other vectors is found.

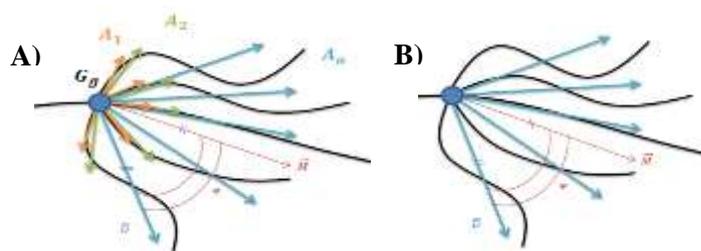


Figure 4. Building the family of vectors along learning trajectories. (A) demonstrates the building of vectors for different fragments of the child and mother trajectories; B) demonstrates final vector construction for calculating the divergence estimation for some selected duration of the learning procedure.

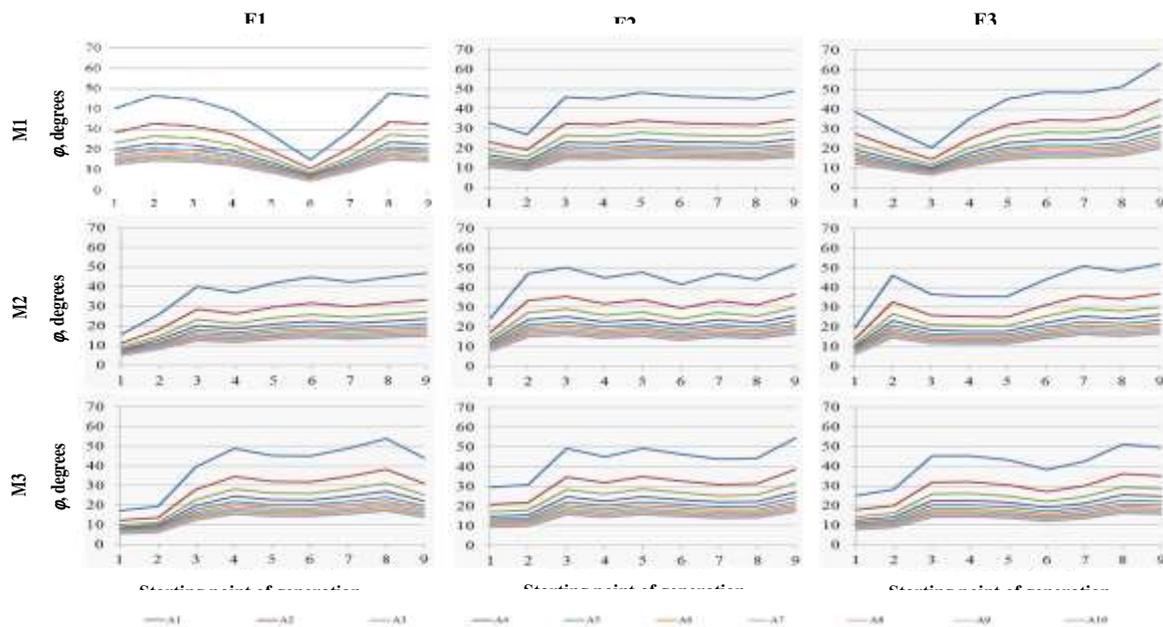


Figure 5. Changes in the estimation of trajectory divergence for three investigated functions (F_i) and for three different initial points (M_i). Different graphs (A_i) correspond to different durations of the estimation learning process, the more i , the longer the learning process.

For any vector D of the selected group of vectors and the central vector of this group M (assuming that the length of D and M are equal), the angle between them is found by the formula

$$\varphi = \arccos \left(1 - \frac{|\vec{D} - \vec{M}|^2}{2|\vec{D}|^2} \right). \quad (4)$$

This parameter characterizes the divergent component of evolutionary movement at a given point on a trajectory and visually (in 3D) can be represented as the degree of disclosure of ‘a vector fan’ of potential trajectories emanating from this point.

This estimation was applied to 9 points along the trajectories under investigation (figure 5). The simulation results showed that the degree of determinism of evolutionary trajectories does not have any significant dependence on the training stage (G_i), on the type of function (F_j) or on the location of the starting point on the landscape (M_k).

4. Discussion

The main result of this work is to demonstrate the inapplicability of the evolutionary concept based on the idea of the single endpoint of evolution [7]. The ability to evolve itself is inextricably linked with the existence of degrees of freedom to avoid local extremes associated, in turn, with the redundancy of the structure of the evolving system over the complexity of the function under selection. Structural redundancy means that the same function can be realized by different structures that in turn signifies the multiplicity of evolutionary outcomes and the principal divergence of evolutionary trajectories. Moreover, each point on an evolutionary trajectory is the point of divergence that was shown by conducted computational experiments.

The obtained results are in good agreement with real experiments with proteins on the evaluation of probability of occurrence of various ANN structures implementing the same function [4], [5]. It was shown that the considered functions of proteins can be realized by different structures and this possibility can be realized in the course of real evolution. Moreover, these structures organize some clusters similar to those obtained in this work and transitions between clusters have tangible probability.

Therefore, the question regarding the distance between trajectories has no definite answer because the distance between trajectories increases on average in the course of evolution. We can discuss the distance between the endpoints of the evolution process which is determined by the size of the corresponding cluster and the distances between clusters.

5. Conclusions

This work is devoted to the conceptual problem and therefore the maximum result we can expect from any heuristic model is generation of some hypothesis about real systems, which can organize experiments and then can be checked in these experiments.

The hypotheses which emerge from the presented heuristic model states that “the divergence of the trajectories is a manifestation of the functional invariance of any evolving structure, and the observed evolutionary changes are some combination of changes in the structure towards the growth of the fitness function and a functional-invariant variation of structure parameters (similar to neutral Kimura mutations [21]), which does not change the level of fitness”. In accordance with obtained results, the existence of functionally invariant variations of structure has to be manifested in experiments via an almost constant level of divergence of evolutionary trajectories at any distance from the evolution endpoints.

Further development of this work is seen in modifying the ANN training algorithm in the direction of greater similarity to the mutation process, and qualitative and quantitative comparison of ANN

trajectories with ones produced by other evolution models and with evolutionary trajectories of real systems.

References

- [1] Lobkovsky A E, Wolf Y I and Koonin E V 2011 Predictability of evolutionary trajectories in fitness landscapes *PLoS computational biology* **7**12 e1002302
- [2] Povolotskaya I S and Kondrashov F A 2010 Sequence space and the ongoing expansion of the protein universe *Nature* **465** 922-6
- [3] Orgogozo V 2015 Replaying the tape of life in the twenty-first century *Interface Focus* **5** 620150057
- [4] Podgornaia A I and Laub M T 2015 Pervasive degeneracy and epistasis in a protein-protein interface *Science* **347** 673-7
- [5] Starr T N, Picton L K and Thornton J W 2017 Alternative evolutionary histories in the sequence space of an ancient protein *Nature* **549** 409-13
- [6] Xue K S 2017 Parallel evolution of influenza across multiple spatiotemporal scales *Elife* **6** e26875
- [7] Lobkovsky A E and Koonin E V 2012 Replaying the tape of life: quantification of the predictability of evolution *Frontiers in genetics* **3** 246
- [8] Dunham E J 2009 Different Evolutionary Trajectories of European Avian-Like and Classical Swine H1N1 Influenza A Viruses *Journal of virology* **83** (11) 5485-94
- [9] Poelwijk F J 2007 Empirical fitness landscapes reveal accessible evolutionary paths *Nature* **445** 383-6
- [10] Von Neumann J and Burks A W 1966 Theory of self-reproducing automata *IEEE Transactions on Neural Networks* **51** 388
- [11] Bernal J D 1965 Molecular structure, biochemical function, and evolution *Theoretical and Mathematical Biology* 96-135
- [12] Eigen M 1971 Selforganization of matter and the evolution of biological macromolecules *Naturwissenschaften* **58** 10 465-523
- [13] Forst C V, Reidys C and Weber J 1995 Evolutionary dynamics and optimization *European Conference on Artificial Life* 128-47
- [14] Schuster P 1996 How does Complexity Arise in Evolution *Complexity* **2** 22-30
- [15] Huynen M, Stadler P and Fontana W 1996 Smoothness within ruggedness: The role of neutrality in adaptation *Proc. Natl. Acad. Sci* **93** 397-401
- [16] Wolpert D H and Macready WG 1997 No Free Lunch Theorems for Optimization *IEEE Transactions on Evolutionary Computation* **1** (1) 67-82
- [17] Bartsev S I and Bartseva O D 2002 Symmetric Structures and Equifinality of Evolution Outcomes in Simple Neural Network Models *Doklady Biochemistry and Biophysics* **386** 235-8
- [18] Bartsev S I 2004 Essence of life and multiformity of its realization: expected signatures of life *Adv.Space Res.* **33**(8) 1313-7
- [19] Bartsev S I and Bartseva O D 2006 Functional Invariant Approach to the Biological System Uniqueness: a Simple Neuronet Model *Doklady Biochemistry and Biophysics* **406** 15-8
- [20] Bartsev S I and Okhonin V A 1991 Variation principle and algorithm of dual functioning: examples and applications *Neurocomputers and attention II* 445-52
- [21] Kimura M 1991 Recent development of the neutral theory viewed from the Wrightian tradition of theoretical population genetics *PNAS* **88** 5969-73