INTERACTION BETWEEN LEARNING AND EVOLUTION IN POPULATIONS OF AUTONOMOUS AGENTS

Vladimir G. Red'ko

Scientific Research Institute for System Analysis, Russian Academy of Science,
Vavilova Str., 44/2, Moscow, 119333, Russia,
vgredko@gmail.com, http://www.niisi.ru/iont/staff/rvg/index_eng.php

Abstract: *The model of interaction between learning and evolution for the evolving population of modeled organisms is designed and investigated. The mechanism of genetic assimilation of the acquired features during the numerous generations of Darwinian evolution is studied. The mechanism of influence of the learning load is analyzed. It is showed that the learning load leads to a significant acceleration of an evolution. The hiding effect is also studied. This effect means that a strong learning inhibits the evolutionary search in some situations.*

Keywords: *Speed and efficiency of evolutionary search, Baldwin effect, genetic assimilation.*

1. INTRODUCTION

In the XIX century, the concepts, suggesting that interaction between learning and evolutionary processes is possible, appeared [1-3]. According to these concepts, learning can contribute significantly to an evolutionary process. This type of influence of learning (or other acquisition of useful features during the life of the organism) on the evolutionary process is often called the Baldwin effect [1]. According to this effect initially acquired features can become inherited during a number of generations. The evolutionary "re-invention" of useful features, initially obtained by means of learning, is often called genetic assimilation [4].

Number of works attempted to model and to analyze interactions between learning and evolution by means of computer simulations [5-10]. In particular, Hinton and Nowlan demonstrated that learning can learning can guide an evolutionary process to find the optimum [7]. Mayley investigated different aspects of interaction between learning and evolution [8] and demonstrated that the hiding effect can take place if the learning is sufficiently strong. The hiding effect means that if learning increases the chances of finding a good phenotype independently on the genome, then learning can also inhibit the evolutionary optimization and genetic assimilation. In addition, the learning load (the cost of learning) was taken into account in [8]. The learning load means that the process of learning has an additional load for the organism and fitness of the organism is reduced under the influence of the load.

The interaction between learning and evolutionary optimization of a neural network control system of autonomous agents was modeled in [10]. The genetic assimilation of the acquired features of agent was observed during several generations of evolution. It was demonstrated that learning can significantly accelerate a process of evolutionary optimization. However, it was difficult to analyze the detailed mechanism of interaction between learning and evolution in these models [10], because these mechanisms were «hidden» in the dynamics of numerous synapse weights of neural networks of agents.

This article develops the mentioned works; it uses works [7, 8] as background. In contrast to [7, 8], the current work uses one of the most clear evolution model, namely, the quasispecies model, proposed by Eigen [11, 12], and quantitative estimations of evolutionary rate and the effectiveness of evolutionary algorithms, obtained in [13, 14]. The quasispecies model considers the process of evolution that is based on the selection and mutations of the genomes of organisms (without crossovers) and describes the main properties of the evolutionary process. The use of models and methods of works [11-14] allows getting a better understanding of the mechanisms of interaction between learning and evolution. In particular, our approach allows analyzing quantitatively the mechanism of genetic assimilation.

2. DESCRIPTION OF THE MODEL

The evolving population of modeled organisms (or individuals) is considered. Similar to [7] we assume that there is a strong correlation between the genotype and the phenotype of the modeled organisms. We assume that the genotype (or the genome) and the phenotype of the organism have the same form, namely, they are chains; symbols of both chains are equal to 0 or 1. The length of these chains is equal to N. For example, similar to [7], we can assume that the genome encodes a model chain DNA, «letters» of which are equal to 0 or 1, and the phenotype is determined by the neural network of organisms, the synaptic weights of the neural network are equal to 0 or 1 too. The initial synaptic weights, received at birth of the organism, are determined by the genome (more precisely, the synaptic weights are equal to the genome symbols). These synaptic weights are changed by means of learning during organism's life.

Similar to the quasispecies model, we assume that each organism has its own genome S_0. The population consists of n organisms, organism's genomes are equal to S_{0k}, $k = 1,..., n$. The organism genome S_{0k} is a chain of symbols, S_{0ki}, $i = 1,..., N$. We also assume that the length of the chains N and the number of organism in population n are large: N, $n \gg 1$. The values N and n don't change in the course of evolution. Symbols S_{0ki} are equal to 0 or 1. We assume that N is so large that only a small part of possible 2^N genomes can be presented in a particular population: $2^N \gg n$. Typical values of N and n in our computer simulations are as follows: $N \sim n \sim 100$.

The evolutionary process consists of a sequence of generations. The new generation is obtained from the old one by selection and mutations. Genomes of organisms of the initial generation are random.

In order to consider learning processes, we introduce two types of sequences: 1) genomes or initial sequence S_{0k} that is received by the organism at its birth, and 2) the current sequence of the organism S_{Tk}.

Organisms inherit the genomes S_{0k} from their parents, these genomes don't change during the organism life and are transmitted (with small mutations) to their descendants. Mutations are random changes of symbols S_{0ki}. The organism receives the genome at its birth, the current sequence S_{Tk} at the birth time moment is equal to the genome: $S_{Tk}(t = 1) = S_{0k}$. The life time of any organism is equal to T. The time is discrete: $t = 1,...,T$. The duration of the generation is equal to T. The sequence of S_{Tk} is modified during the organism life by means of learning. The current sequence S_{Tk} determines the organism's phenotype.

As descendants of organisms obtain just genomes S_{0k} that organisms received from their parents and not sequences S_{Tk} that are optimized by learning, the evolutionary process has the Darwinian character.

It is assumed that there is an optimal sequence of S_m (components of which are also equal to 0 or 1), which is searched for in the processes of evolution and learning. At computer simulation, the sequence S_m was set to be the random one.

Learning is performed by the following method of trial and error. Every time moment t each symbol of the sequence S_{Tk} is randomly changed to 0 or 1, and if this new symbol coincides with the corresponding symbol of the optimal sequence S_m, then this symbol is fixed in the S_{Tk}; otherwise, the old symbol of the sequence S_{Tk} is restored. So, during learning, the current sequence S_{Tk} moves towards the optimal sequence S_m.

It should be noted that if we consider symbols of sequences S_{Tk} as synaptic weights of the neural network, then the learning process has a simple meaning: learning is searching for optimal weights of the synapses.

At the end of the generation, the selection of individuals in accordance with their fitness takes place. The fitness is determined by the sequence S_{Tk} at the time moment $t = T$. We denote this sequence of S_{Fk}, i.e. we set $S_{Fk} = S_{Tk}(t = T)$. The fitness of the organism S_k is determined by the Hamming distance $\rho = \rho(S_{Fk}, S_m)$ between the sequences S_{Fk} and S_m:

$$f(S_k) = \exp[-\beta\rho(S_{Fk}, S_m)] + \varepsilon, \qquad (1)$$

where β is the positive parameter, that characterizes the intensity of selection, $0 < \varepsilon \ll 1$. The role of the parameter ε in (1) can be considered as the influence of random factors of the environment on the fitness of individuals.

The selection of the individuals into a new generation is made by the well-known method of the fitness proportionate selection (or the roulette wheel selection). In this method individuals are selected into a new generation probabilistically. The choice of an individual into the next generation takes place n times, so the number of individuals in the population in all generations is equal to n. At any choice, the probability of the selection of a particular individual is proportional to its fitness.

Thus, individuals are selected at the end of a generation in accordance with their phenotype codes $S_{Fk} = S_{Tk}(t = T)$, i.e. in accordance with the final result of learning, whereas initial genomes S_{0k} (modified by small mutations) are transmitted from parents to descendants.

In addition, similar to the work [8], we take into account the learning load (or the cost of learning), namely, we consider the fact that the learning

process has a certain burden on the individual and fitness of the individual may be reduced under the influence of the load. For this purpose we use the modified fitness of individuals:

$$f_m(\mathbf{S}_k) = \exp(-\alpha d) \{\exp[-\beta\rho(\mathbf{S}_{Fk},\mathbf{S}_m)] + \varepsilon\} , \qquad (2)$$

where α is the positive parameter, which takes into account the learning load, $d = \rho(\mathbf{S}_{0k},\mathbf{S}_{Fk})$ is the Hamming distance between the initial $\mathbf{S}_{0k}$ and final sequence $\mathbf{S}_{Fk}$ of the individual, i.e. the value that characterizes the intensity of the whole learning process of the individual during its life. The factor $\exp(-\alpha d)$ decreases the fitness of an individual, this decrease clearly depend on the change of the current sequence $\mathbf{S}_{Tk}$ at the learning process.

It should be noted that since initial sequences $\mathbf{S}_{0k}$ of the individuals in the initial population are random, the average Hamming distance between these sequences and the optimal one $\mathbf{S}_m$ is approximately equal to $N/2$. The sequences $\mathbf{S}_k$ should overcome this distance by means of learning and evolution in order to reach $\mathbf{S}_m$.

3. RESULTS OF SIMULATION

3.1. SCHEME AND PARAMETERS OF SIMULATION

Two modes of operation of the model are consider below: 1) evolution combined with learning, as described above, 2) "pure evolution", that is evolution without learning, in this case, the learning doesn't take place and it is believed that $\mathbf{S}_{Tk} = \mathbf{S}_{0k}$. In addition, the influence of the learning load is analyzed; in this case, the fitness of an individual is calculated according to (2). Analysis of the model was carried out by means of computer simulation.

The parameters of the model at simulation are chosen in such manner that the evolutionary search is effective; the experience of the work [13] for the case of pure evolution is used at this choice. The fitness of the individuals in [13] was determined analogously to the expression (1), only the influence of random factors wasn't taken into account (formally this means that the value ε was equal to 0).

The choice of parameters of simulation is as follows. We believe that the length of the chain is rather large: $N = 100$. We also set $\beta = 1$, this corresponds to a sufficiently high intensity of selection, so the selection time is small, the time of evolutionary search is determined mainly by mutations. The intensity of mutations must not be too large; in order to remove the possibility of mutation losses of already found good individuals. However, the intensity of mutations must not be too small, in order to ensure sufficiently large efficiency of mutational search during evolutionary

optimization. Taking this into account, we believe that the probability to substitute any symbol in the sequence $\mathbf{S}_{0k}$ at mutations in one generation is $p_m = N^{-1} = 0.01$. At this mutation intensity p_m approximately one symbol in the genome of any individual is replaced at one generation, i.e. during one generation of the Hamming distance ρ between sequences of individuals $\mathbf{S}_k$ in population and the optimal sequence $\mathbf{S}_m$ changes on average by 1 by means of mutations. Selection leads to a decrease of this distance ρ. Since the intensity of selection is large, and the Hamming distance between sequences of the initial population and the optimal sequence $\mathbf{S}_m$ is of the order of N, the whole process of evolution will take approximately $G_T \sim N$ generations. Such an estimation of the rate of evolution is true, if the population size is large enough and the fluctuation effects and neutral selection of individuals (that is selection independent on fitness of individuals) can be neglected. To satisfy this condition, it is enough to require that the characteristic time of the neutral selection (which is of the order of the population size n [13, 15]), should be greater or of the order of G_T, so we believe that $n = G_T = N$.

Thus, the parameters of simulation in accordance with the experience of the work [13] are chosen as follows: $N = 100$, $\beta = 1$, $p_m = N^{-1} = 0.01$, $n = G_T = N = 100$.

In the current model we also believe that the probability of a random replacement of the symbols during learning p_l is rather large: $p_l \sim 1$, the number of time moments during the generation T is equal to 2 (choice of such parameters p_l and T means that learning is rather fast), the parameter ε is small: $\varepsilon = 10^{-6}$.

The results of simulation are averaged over 1000 or 10000 calculations corresponding to different random number generators. The results of simulation are described below.

3.2. COMPARISON OF REGIMES OF PURE EVOLUTION AND EVOLUTION COMBINED WITH LEARNING

Fig. 1 shows the dependence of the average of the Hamming distance $\rho = \rho(\mathbf{S}_k,\mathbf{S}_m)$ between the sequences $\mathbf{S}_k$ of the individuals in the population and the optimal sequence $\mathbf{S}_m$ at the beginning of generations (i.e. in this case $\rho(\mathbf{S}_k,\mathbf{S}_m) = \rho(\mathbf{S}_{0k},\mathbf{S}_m)$) on the generation number G . Curve 1 characterizes the regime of evolution combined with learning; curve 2 characterizes the regime of pure evolution. The dependences are averaged for all individuals of population and for 1000 calculations. Fitness of individuals is determined by the expression (1). We can see that pure evolution without learning (curve 2) doesn't optimize individuals $\mathbf{S}_k$ at all even at

small values ε; whereas evolution combined with learning (curve 1) obviously ensures the movement towards the optimal individual S_m.
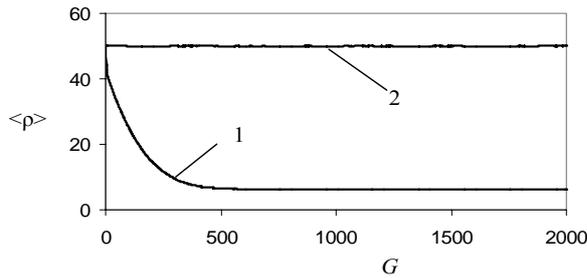

Fig. 1 – The dependence of <ρ> on the generation number G. Curve 1 characterizes the regime of evolution combined with learning; curve 2 characterizes the regime of pure evolution

To understand, why the pure evolution doesn't ensure a decrease the value ρ, let's estimate the value of fitness (1) in the original population. Individuals of the initial population S_{0k} are far from the optimal one S_m: the Hamming distance $\rho = \rho(S_{0k}, S_m)$ is of the order of $N/2 = 50$, therefore, $\exp(-\rho) \sim 10^{-22}$ and $\exp(-\rho) \ll \varepsilon$. This means that all the individuals of the population have approximately the same value of the fitness $f(S_k) \approx \varepsilon$. Consequently, a selection of the individuals doesn't occur in the case of the pure evolution. Thus, the movement towards S_m occurs only in the presence of learning; this movement leads to the decrease of ρ. A similar influence of learning on evolutionary optimization (though in another context) was described in the work [7].

Let's consider the effect of the acceleration of the evolutionary process by learning (curve 1 in Fig. 1). Analysis of the results of simulations shows that the gradual decrease in the values $\rho = \rho(S_k, S_m)$ occurs as follows. During learning, the distribution of individuals $n(\rho)$ on the value ρ in the population is shifted towards smaller ρ, so the values $\rho = \rho(S_{Fk}, S_m)$ becomes small enough, then $\exp(-\rho(S_{Fk}, S_m))$ is of the order of ε. Consequently, different individuals in the population in accordance with (1) have different fitness; so individuals with small values $\rho(S_{Fk}, S_m)$ are selected into the population of the next generation. It is intuitively clear that the genomes of S_{0k} of selected individuals should be rather close to the final sequences S_{Fk} (obtained as a result of the learning) of these individuals. Thus, the result of selection is the selection of individuals, which genomes are also moving to the optimal sequence S_m. Therefore, the value ρ in the new population decreases.

The described mechanism is characterized by Fig. 2, which shows the distribution the $n(\rho)$ on ρ in the population at different moments of the first generation. Curve 1 shows the distribution of the $\rho = \rho(S_{0k}, S_m)$ for the genomes of individuals at the beginning of the generation. Curve 2 shows the

distribution of the $\rho = \rho(S_{Fk}, S_m)$ for individuals after the learning, but before selection. Curve 3 shows the distribution of the $\rho = \rho(S_{Fk}, S_m)$ for individuals, selected in accordance with the fitness (1). Curve 4 shows the distribution of the $\rho = \rho(S_{0k}, S_m)$ for the genomes of selected individuals at the end of the generation. The genomes of selected individuals S_{0k} are sufficiently close to the sequences of trained and selected individuals S_{Fk}, therefore the distribution of the $\rho = \rho(S_{0k}, S_m)$ for genomes (curve 4) moves towards the distribution for finite sequences S_{fk} (curve 3). Finally, after the selection at the end of the generation, the distribution of the genomes of ρ (curve 4) is formed; this distribution is closer to the distribution, which is represented by curve 3, than the initial distribution of genomes (curve 1). Similar displacement of the distribution of $n(\rho)$ towards smaller values ρ takes place in the next generations.
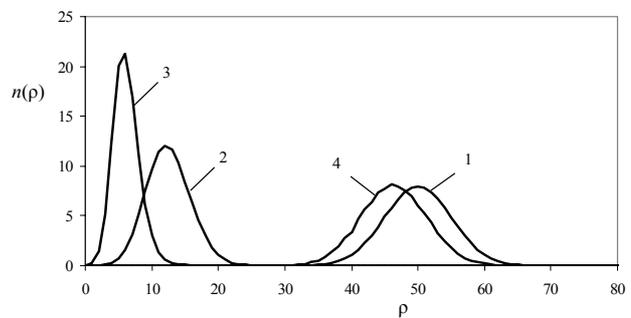

Fig. 2 – The distribution n(ρ) in the first generation of evolution: curve 1 is the distribution of ρ = ρ(S0k,Sm) for the original genomes before learning, curve 2 is the distribution of ρ = ρ(SFk,Sm) for individuals after the learning, but before the selection, curve 3 is the distribution of ρ = ρ(SFk,Sm) for selected individuals, curve 4 is the distribution of ρ = ρ(S0k,Sm) for the genomes of selected individuals at the end of the generation (results are averaged on 10,000 calculations)

Such displacement reveals the mechanism of reduction of <ρ> in the presence of learning: the selection leads to the genomes of individuals S_{0k}, which are closer to sequences of learned and selected individuals S_{Fk}, than the initial genomes of individuals at the beginning of the generation. As a result, a transition from curve 1 to curve 4, i.e. the decrease in the values ρ, takes place during a generation.

It should be underlined that the decrease of values ρ at learning should be sufficiently large in order to ensure small role of the parameter ε and significant difference of the fitness (1) of individuals, and therefore, the effective selection of individuals with small values $\rho(S_{Fk}, S_m)$. This selection corresponds to the essential decrease of the values of ρ at transition from curve 2 to curve 3 in Fig. 2. It is clear that in order to guarantee the effective operation this mechanism, learning should

be enough strong. The other role of a strong learning is characterized in the next subsection.

The described results show that learning can lead to the effective genetic assimilation and to the radical acceleration of the evolutionary search.

3.3. HIDING EFFECT

However, a strong learning can not only accelerate the evolutionary search, but it can prevent to find the optimal genome. Curve 1 in Fig. 1 shows that at large G, the decrease of $<\rho> = <\rho(\mathbf{S}_{0k}, \mathbf{S}_m)>$ is limited: the final value of the $<\rho>$ remains quite large, the asymptotic value of the $<\rho>$ is approximately equal to 6.2. This is due to the fact that at large G (G ~1000) the strong learning ($p_l = 1$, $T = 2$) results in finding the optimal sequence $\mathbf{S}_m$ independently on the genome $\mathbf{S}_{0k}$. Therefore, at the final stages of evolution the genomes $\mathbf{S}_{0k}$ don't move towards the optimum $\mathbf{S}_m$. So, the hiding effect [8] is observed.

Thus, the mechanism of the hiding effect is analyzed. This effect means that the strong leaning prevents the evolutionary optimization, because it increases the chances of finding a good phenotype independently on the genome of the individual.

3.4. INFLUENCE OF LEARNING LOAD ON MODELED PROCESSES

We also analyzed the influence of the learning load on the modeled processes. For this case, fitness is determined by the expression (2). The simulation is performed for the mentioned parameters ($N = n =100$, $\beta = 1$, $p_m = 0.01$, $p_l = 1$, $T = 2$, $\varepsilon = 10^{-6}$), the value α is equal to 1. The simulation results are represented by Fig. 3, 4. Fig. 3 shows the dependence of the average distance $<\rho>$ between sequences $\mathbf{S}_k$ and the optimal sequence of $\mathbf{S}_m$ on the generation number G. Fig. 4 shows the dynamics of the distribution $n(\rho)$ of values ρ at different moments of the first generation of the evolutionary process.
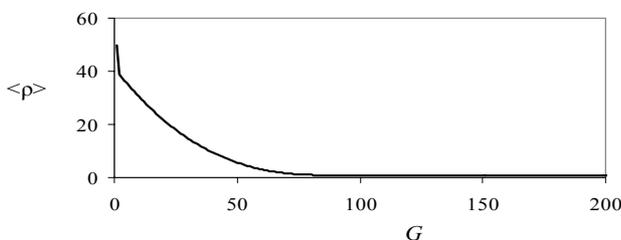

Fig. 3 – The dependence of the value $<\rho>$ on the generation number G; the fitness of individuals is determined by the expression (2); the decrease of values $<\rho>$ is much faster than that of in Fig. 1 (results are averaged on 1000 calculations)
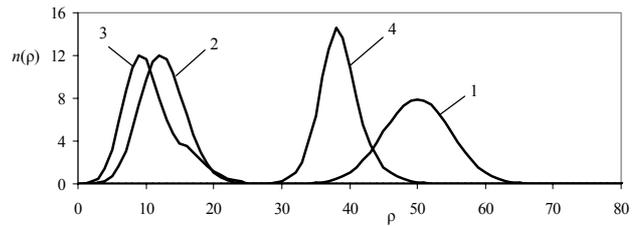

Fig. 4 – The distribution n(ρ) in the first generation of evolution; the fitness of individuals is determined by the expression (2): curve 1 is the distribution of ρ = ρ(S0k,Sm) for the original genomes before learning, curve 2 is the distribution of ρ = ρ(SFk,Sm) for individuals after the learning, but before the selection, curve 3 is the distribution of ρ = ρ(SFk,Sm) for selected individuals, curve 4 is the distribution of ρ = ρ(S0k,Sm) for the genomes of selected individuals at the end of the generation; the displacement of the distributions to smaller values ρ is significantly faster than in Fig. 2 (results are averaged on 10,000 calculations)

It should be underlined that the genetic assimilation for cases of the fitness, determined by the expression (1) and the expression (2), has the same nature. In both cases, the genomes of selected individuals $\mathbf{S}_{0k}$ approach to sequences $\mathbf{S}_{Fk}$ of trained and selected individuals. That is in both Fig. 2 and Fig. 4 the curve 4 moves towards the curve 3. A significant difference consists only in the fact that the learning load makes this movement more evident and more effective. Thus, the learning load leads to more effective selection of individuals with the genomes of $\mathbf{S}_{0k}$, which are close to $\mathbf{S}_m$, and the evolution process is significantly accelerated.

4. CONCLUSION

Thus, the model of interaction between learning and evolutionary optimization has been constructed and investigated.

The mechanism of the genetic assimilation is studied in details. The genetic assimilation can be described as follows: 1) learning and selection shift the distribution of individuals towards the optimum; 2) genomes of selected individuals also move towards the optimum. The mechanism of the genetic assimilation is illustrated by Fig. 2. It is shown that the genetic assimilation can lead to a radical acceleration of evolutionary processes.

The mechanism of the hiding effect is analyzed. This effect means that strong learning inhibits the evolutionary search of the optimal sequence, if this learning increases the chances of finding a good phenotype regardless of the genome of the individual.

The influence of the learning load on the evolutionary processes is studied. It is shown that the learning load leads to effective genetic

assimilation and to a considerable acceleration of evolution.

Future plans of our research include investigations of cognitive features of autonomous agents. These agents can be optimized by means of learning and evolution.

5. ACKNOWLEDGMENTS

6. REFERENCES

[1] J.M. Baldwin, A new factor in evolution, *American Naturalist*, (30) (1896), pp. 441-451.

[2] C.L. Morgan, On modification and variation, *Science*, (4) (1896), pp 733-740.

[3] H.F. Osborn, Ontogenetic and phylogenetic variation, *Science*, (4) (1896), pp. 786-789.

[4] C.H. Waddington, Canalization of development and inheritance of acquired characters, *Nature*, (150) (1942), pp. 563-565.

[5] *Adaptive Organisms in Evolving Populations: Models and Algorithms.* Eds. Belew R.K. and Mitchell M. Massachusetts: Addison-Wesley, 1996.

[6] *Evolution, Learning, and Instinct: 100 Years of the Baldwin Effect.* Eds. Turney P., Whitley D., Anderson R. Special Issue of Evolutionary Computation on the Baldwin Effect, (4) 3 (1996).

[7] G.E. Hinton, S.J. Nowlan, How learning can guide evolution, *Complex Systems*, (1) (1987), pp. 495-502.

[8] G. Mayley, Guiding or hiding: Explorations into the effects of learning on the rate of evolution, In: *Proceedings of the Fourth European Conference on Artificial Life (ECAL 97).* Eds. Husbands P. and Harvey I. Cambridge, Massachusetts: MIT Press,1997, pp. 135-144.

[9] D. Ackley, M. Littman, Interactions between learning and evolution, In: *Artificial Life II: Proceedings of the Second Artificial Life Workshop.* Eds. Langton C. G., Taylor C., Farmer J. D., Rasmussen S. Redwood City CA: Addison-Wesley, 1992, pp. 487-509.

[10] V.G. Red'ko, O.P. Mosalov, D.V. Prokhorov, A model of evolution and learning, *Neural Networks*, (18) 5-6 (2005), pp. 738-745.

[11] M. Eigen, Selforganization of matter and the evolution of biological macromolecules, *Naturwissenschaften*, (58) 10 (1971), pp. 465-523.

[12] M. Eigen, P. Schuster, *The hypercycle: A principle of natural self-organization.* Springer Verlag: Berlin etc, 1979.

[13] V.G. Red'ko, Yu.R. Tsoy, Estimation of the efficiency of evolution algorithms, *Doklady Mathematics,* (72) 2 (2005), pp. 810-813.

[14] V.G. Red'ko, Yu.R. Tsoy, Efficiency of evolutionary search in quasispecies model, *Fuzzy Systems and Soft Computing*, (1) 1 (2006).

[15] M. Kimura, *The Neutral Theory of Molecular Evolution*, Cambridge University Press, 1983.

Red'ko Vladimir Georgie-vich, *Deputy Director for Research of Center of Optical Neural Technologies, Scientific Research Institute for System Analysis, Russian Academy of Sciences. V.G. Red'ko gradua-ted from the Moscow Institute of Physics and Technology in 1971. He is the doctor of sciences (physics and mathematics). He is the author of more than 150 scientific publications, including two monographs. His scientific interests includes: the problem of origin of human intelligence, cognitive evolution, models of adaptive behavior, neuroinformatics.*