Evolutionary Theories and Genetic Algorithms

Hristo I. Toshev[1], Stefan L. Koynov[2] and Chavdar D. Korsemov[3]

Abstract: - **Some of the basic theories of evolution are analyzed in this paper and the possibilities for different applications of their ideas are introduced. Of special interest are the characteristics and the peculiarities of genetic algorithms (GA) that distinguish them from traditional optimization methods and procedures for searching, the stages of preliminary preparation for solving practical problems by GA included.**

Keywords – **genetic algorithms, evolution, selection, mutation, optimization**

I. INTRODUCTION

The word *evolution* (*evolutio* in Latin) means *change*, *development* (slow and gradual). Evolution is usually the opposite of revolution which concerns fast and vast changes that sometimes lead to unpredictable results. At the end of XVIII century Bonnet, a Swiss scholar, uses the term evolution for the first time in biology in the context of a slow and gradual qualitative and quantitative change of the object when each state of the object must be at a higher level of development and organization compared to the previous one.

Modern time is marked by a great number of versions of different concepts for evolution [1], [2], [7], [12]. The unifying factor in their development is the type of changeability that they accept as a basis of evolution, a determinate direction of adaptability or indeterminate when adaptability is just random.

Development of theories of evolution lead to big chances for development of new scientific trends.

Evolution in biology is defined by hereditary changeability, struggle for existence, natural and artificial selection [1].

During early 40-s in 20-th century, in classical genetics D. Morgan postulated the following evolutionary approaches: (i) genes (elementary units of hereditary information in chromosomes) are mutable, they mutate and mutations of particular chromosomes lead to changes of single elementary features; (ii) all features of the organism are controlled by genes.

Later based on elaborations of M. Watson and F. Crick a new scientific direction appeared, molecular genetics. Genetic code turns out to be common for all natural systems on our planet and biological specifics of living creatures are inherited according to definite rules (discovered by Mendel).

New scientific branches of knowledge originated, e.g. bionics (*bion* in Greek means *element of life*) that joints together biology, genetics and technique aimed at solving engineering-technical problems on the basis of genetics and analysis of vitality of organisms.

[1] Hristo I. Toshev, [2] Stefan L. Koynov and [3] Chavdar D. Korsemov are with the Institute of Information Technologies, Bulgarian Academy of Sciences, Acad. G. Bonchev str., bl. 29A, 1113 Sofia, Bulgaria, E-mail: toshev@iinf.bas.bg, slk@iinf.bas.bg, chkorsemov@iinf.bas.bg

At the end of the 50s scientists from different countries [7], [8], [10] explored in detail evolutionary systems and independently came to the conclusion that they could use the theory of evolution as an instrument for optimization in the process of solution for problems of different nature with the main goal creating a population of eventual solutions using some of the most characteristic peculiarities of nature – heredity, changeability, selection and so on.

Genetic algorithms are a method for search based on the selection of the best species in the population in analogy to the theory of evolution of Ch. Darwin [7].

Their origin is based on the model of biological evolution and the methods of random search. From the bibliographical sources [3], [7], [8], [10] it is evident that the random search appeared as a realization of the simplest evolutionary model when the random mutations are modelled during random phases of searching the optimal solution and the selection is modeled as "removal" of the unfeasible versions.

The main goal of GA-s is twofold:

- abstract and formal explanation of the adaptation
- modelling natural evolutionary processes for efficient solution of determined class of optimization and other problems.

During the last years a new paradigm is applied to solve optimization problems GA-based and modifications of GA. GA realize searching a balance between efficiency and quality of solutions at the expense of selecting the strongest alternative solution from undetermined and fuzzy solutions [12].

II. EVOLUTIONARY THEORIES

2.1. Lamarck's Theory

Lamarck in the beginning of 19-th century (in 1809) [2] presented the first group of concepts and hypotheses of evolution. He assumed that all living creatures adapt expediently to the environmental conditions. He accepted that the reasons for evolution are: the trend towards a progress, a development from simpler to more sophisticated (gradation) and also a modification and an adaptation of organisms to their surroundings.

These changes are provoked directly by the external conditions thus leading to an improvement of the organs and also to an inheritance of the acquired features. These statements are considered a basis of the theological theory of evolution; according to it the ability of creatures to adapt to their environment is their inborn property.

Lamarck explains one of the peculiarities of the organic world with the adaptability. The progressive evolution, the origin of new species (more sophisticated and more perfect) he explains with laws of gradation, the tendency of living creatures to complicate their own structure:

- first law. An improvement (drill) of organs leads to their progressive development. The absence of improvement of organs leads to a reduction.
- second law. The results of improvements and their absence for long periods of stimulations are accepted by the organism as new qualities and consequently they are handed down from one generation to another depending on the influence of the environment.

Though that biologists reject Lamarck's theory, it is a powerful concept for the artificial evolution that is used in science and technique because it allows local searching in cases of complex practical problems with big dimensionalities for a research of local domains of the population.

2.2. Charles Darwin's Theory

Charles Darwin is the first scientist who determined a common principle in living nature; this principle is the natural selection [1]. There are two aspects in his theory: the materials of evolution, and its factors and rootive powers. Strictly speaking, the rootive power of evolution is the natural selection. Selection is used to achieve the best results of development of species. As a science it is created by Charles Darwin; he chose three forms:

- natural selection. It provokes changes connected with an adaptation of the population to the new conditions.
- unconscious selection. The best species (exemplars) are saved in the population.
- methodic selection. The changes in the population are purposeful.

The rootive powers of evolution are the following:

- indefinite changeability, i.e. a hereditarily determined variety of organisms of every population;
- struggle for existence when less adapted organisms are eliminated from reproduction;
- natural selection. The exemplars that are better adapted survive. As a result, useful hereditary changes are accumulated and integrated thus producing new adaptations.

Based on his theory, Darwin proved that the main rootive power of evolution is the selection of the best species. Their realization means that the following conditions must be observed: a correct selection of the original material, a strict determination of the goal, an implementation of the selection in sufficiently wide ranges and a quick removal of the rejected material for the selection of a single basic feature.

2.3. Fries's Theory

In his research [5], [6] he does not take into consideration the integrating role of the natural selection. Instead he bases his platform on the basis of manifestation of mutations as direct factors for the formation of species.

He develops the ideas of saltacionism where the basic way for the origin of species are sudden jumps (quick changes) without any preliminary accumulation of qualitative changes in the evolutionary process. This type of evolutionary mechanism is called an evolutionary failure. It is manifested once per several thousand generations. The basic idea lies in the imported global changes in the contents of the genes on catastrophic occasions. The origin of species in moments of sudden leaps is a real process.

Similar research in this direction was performed also by scientists like Jan Potszki, U. Batton, S. Goldschmidt, W. Kordlam, etc.

2.4. Popper's Theory

The concept of K. Popper [4] belongs to the basic types of discussed theories. He presents an evolutionary epistemology interpreted as a triad deductivism – selection – critical removal of errors.

The exposure is in the form of theses the basic of which are the following:

- all evolutionary systems solve definite problems;
- problems are always solved using the method of trials and errors;
- errors are removed by complete removing of unsuitable species or by a modification;
- the population uses a mechanism that is elaborated in the process of evolution of the species;
- the population is a basic factor of the evolutionary species to which it belongs. It is a trial solution analyzed in the evolutionary process that chooses the environment and its transformation.

The scheme of the evolutionary sequence of events is the following:

$$F1 \rightarrow TS \rightarrow EE \rightarrow F2$$

where F1 is the source problem, TS is the set of trial solutions, EE is the removal of the errors and F2 is the basic problem.

The difference from Darwin's theory with its single problem – survival of the strongest – and the case with Popper is the presence of other problems with the latter: reproduction, removal of the extra generation and so on; besides here the difference between F1 and F2 is important because it explains the creative intellectual evolution.

According to [4] the information does not penetrate in the system, rather the scientists explore the environment and they actively receive information from it.

Briefly said, according to Popper the evolution is a process of random changeability and the selective memory is in the basis of increase of knowledge and growth of adaptability to the surroundings. Three main parts are evident in this process (triads): changeability mechanism (that joints the selection process), saving and spreading of selected versions.

2.5. Dubinin's Theory

Dubinin's theory [5] has united some of the basic concepts from the theories of Darwin, Lamarck, de Fries and the principles of genetics of populations. The basic concepts are the following:

- evolution is impossible without an adaptation of the organisms to the conditions of the environment. The factor formulating the adaptivity is the natural selection; it uses random mutations and recombinations;
- natural selection based on a transform of genetics of populations creates complex genetic systems. Their modifycations are strengthened by a stabilizing factor;
- hereditary changeability in populations has a massive character. The origin of random mutations is inherent only to single exemplars;

- most adapted exemplars produce a great number of successors;
- species originate by an evolution of populations.
- the contradictions between the random character of hereditary changeability and the requirements of selection define the uniqueness of the types of genetic systems.

Processes that comprise the results from realizations of Dubinin's evolutionary theory include:

- random variations in the size of separate populations, periodic fluctuations, unsteady changes and fluctuations in the size of the population that are determined by the migration process.

N. Dubinin separates four basic forms of realization of the inner unity of the evolutionary process:

- microevolution (processes of intraspecies evolution);
- phase of an increasing evolutionary improvement;
- turning points in evolution;
- evolutions of basic specifics in the organization of the natural system.

III. GENETIC ALGORITHMS

Genetic algorithms (GA) are a new research domain. They originated as a result from the research of D. Holland and colleagues in the University of Michigan, USA. GA that he described borrow their terminology to a great extent from natural genetics. They appeared as a combination of models of biological evolution and also of methods of random search.

We shall present some concepts and definitions from the theories of GA. All GA function on the basis of some initial information; the population of alternative solutions P covers this information. The population $P = \{ p_1 , p_2 ,\dots p_i , \dots, p_{Np} \}$ presents a set of elements p_i. Here N_p is the size of the population. Every element of this population p_1 as a rule presents one or several chromosomes. Chromosomes consist of genes. The positions of genes in the chromosome are called *locii* (*locus* per a single chromosome), i.e. the gene is a subelement (an element in the chromosome), *locus* is position in the chromosome, *allel* is gene's meaning.

Genes may have numeric and functional values. Usually these numeric values are from some alphabet. The genetic material of the elements usually is coded in the binary alphabet {0,1} thought that it is possible to use alphabetic and also decimal, etc. alphabets.

The elements in GA are usually named *parents*. The parents are selected from the population on the basis of predefined rules that are mixed (cross-fertilized) to produce *children* (*successors*). The *generation* (process of realization of a single iteration of the algorithm) is called an *offspring*.

The evolution of the population according to [8], [9] is a succession of generations where the chromosomes change their values in such manner that every new generation adapts to its environment in the best way.

Each element in the population possesses some definite level of quality that is marked by the value of the *goal function* (GF) (sometimes referenced to as *function of utility* or *fitness function*). GF is used in GA to compare solutions and to choose the best one of them. The basic problem of genetic algorithms is the optimization of GF. With other words GA analyze populations of chromosomes that are combinations of elements of some set and after that they optimize GF estimating every chromosome. GA manipulate populations of chromosomes based on the mechanism of natural evolution.

In simple GA created by Goldberg [8] the realization is carried out by three basic genetic operators; selection, crossover and mutation [8].

There exist plenty of different types of selection that we can conditionally divide into probabilistic, determined and combined. We must note that explorers of GA use more and more combined selections with a predefined knowledge about the problem to be solved.

Crossover methods are: simple (single point), paired, sequential, serial, partially corresponding, cyclic, universal, etc. It is important to note that the search of the optimal crossover operator is still going on.

Mutation operators are necessary to avoid losses of important genetic material. More important mutation operators are: single point, paired, of inversions, of translocation, of segregation, of removal and adding.

We can show more precisely the real necessity to use GA-s if we try to present some of their most characteristic peculiarities that distinguish them from traditional optimization methods for search and learning [10], [12], [15]:

- GA operate basically not with the problem parameters, but with a coded set of parameters;
- they realize the search not by improving the solution, but by using several alternatives of the defined set of solutions;
- they use a goal function, not different increments for evaluation of quality of the accepted solutions;
- the accepted rules for analysis of the optimization problem are not determined, but probabilistic.

In operational mode, the GA chooses a set of natural parameters of the optimization problem and it codes them into a sequence with a finite length in some alphabet. The GA operates until a given number of iterations is fulfilled (algorithmic iterations) or during some iteration a solution with a definite quality is obtained or a local optimum is found, i.e. a premature convergence is observed and it is impossible to find an exit from this state [11], [13], [14].

GA-s provide a series of advantages for solving problems in practice. One advantage is the adaptation to the changing environment. In real life the problem that is postulated for solution can undergo vast modifications during the process of its solving. The usage of traditional methods requires that all calculations must be performed from scratch and this leads to big losses of machine time. In the case of evolutionary approaches the population can be analyzed, supplemented and modified in accordance with the modified conditions. For this reason the complete selection is an option, not obligation.

Another important peculiarity in the process of solving practical problems supported by GA is the necessity to execute some preliminary stages:

- *Choose a way for solution presentation* (stage 1). The chosen structure must allow coding of all possible solutions and it also must admit and ensure their estimate [11].
- *Choose random operators for posterity generation* (stage 2). Besides the usage of the two basic types of reproduction,

sexual and asexual (cloning), it is possible to use also operations that do not exist in real nature: for example, the usage of material from three and even more parents that is accompanied by voting for the choice of parents. There are no limitations for the usage of different operators and also it is senseless to copy laws of nature and their limitations mechanically and blindly [3], [7,], [8], [11];

- *Define DM rules for creation and selection of posterities* (stage 3). The simplest choice is the case to accept only the best solutions and to ignore the rest of the decisions. Very often this rule turns out to be little effective and that good decisions may come also from bad ones, i.e. not just from the best ones. Therefore this leads to the logical assumption that the possibility for choice of the best decision must be the most important principle at this stage;

- *Create an initial population* (stage 4). If there is a lack of knowledge for the considered problem then solutions are accessible according to some random choice from the set of possible solutions. This points to a process of generation of random problems when every single problem itself is a definite solution. On the other hand during the process of creation of the primary population it is possible to use data received during the experiment of solving the same problem with other GA-s. If these solutions are really valuable then they will outlast and produce a posterity if they do not perish together with other weak individuals meanwhile [3], [8].

V. Conclusions

The performed analysis of development of some basic evolutionary theories forms the image of a continuous ascending progress of science in this trend. It includes wide possibilities for applications of ideas for evolutionary approaches in cases of solving vast ranges of problems. In order to provide a rough idea of the sort of applications tha are being tackled in the current literature, we will classify the applications in three large groups [17]: engineering, industrial and scientific. A representative sample of engineering applications of the following: electrical engineering, robotics and control [20], structural engineering, hydraulic engineering.

Industrial applications are the following: design and manufacture, scheduling [18], management.

Finally, we have a variety of scientific applications: chemistry [21], physics, medicine, computer science.

GA are a new and effective optimization methodology that is used to solve different optimization problems based on an analogy of natural processes of selection and genetic transformations taking place in nature.

The strong interest for using GA in so many different disciplines reinforces the idea of the multi-objective nature of many real world problems. However some applications domains have received relatively little attention from researchers are represent areas of opportunity. For example [19]: cellular automata, pattern recognition, data mining, bioinformatics and financial applications.

References

[1] Ch.Darwin, *"The Origin of Species by Natural Selection*", Works, V. 3, Moscow-Leningrad, Academia, 1939 (in Russian).

[2] J.-B.Lamarck, "*Philosophy of Zoology"*, V. 1, 2, Moscow-Leningrad, Academia, 1939 (in Russian).

[3] V.V.Emelianov, V.M.Kureichik, V.V.Kureichik, *"Theory and Practice of Evolutionary Modelling"*, Moscow, 2003 (in Russian)л

[4] Evolutionary Epistemology and Logic of Social Sciences: Karl Popper and His Critics // Compilers D.G.Lahuti, N.V. Sadovsky, K.Z.Finn, Moscow, Editorial URSS, 2000 (in Russian).

[5] N.P.Dubinin, *"Selected Works"*. V. 1, "*Problems of the Gene and Evolution"*, Moscow, Nauka, 2000 (in Russian).

[6] E.Meyer, "*Populations, Species and Evolution"*, Moscow, Mir, 1974 (in Russian).

[7] J.H.Holland, *"Adaptation in Natural and Artificial Systems"*, the MIT Press, 1992.

[8] D.L.Goldberg, *"Genetic Algorithms in Search, Optimization and Machine Learning"*, Addison-Wesley, 1989.

[9] L.Davis (Ed.), Handbook of Genetic Algorithms, N.Y., Van Nostrand Reinhold, 1991.

[10] M.Mitchell, *"An Introduction to Genetic Algorithms"*, MA: MIT Press, 1996.

[11] J.R.Koza, *"Genetic Programming 2"*, Cambridge MA: MIT Press, 1994.

[12] D.Goldberg, *"Web Courses"*, Available at: http://www.engr.uiuc.edu/OCCE, 2000.

[13] C.M.Fonseca, P.J.Fleming, Genetic Algorithm for Multiobjective Optimization: Formulation, Discussion and Generalization. In: *Proceedings of the Third International Conference on Genetic Algorithms* (S. Forest, Ed.) Morgan Kaufmann Publishers, pp. 416-423, 1993.

[14] A.Jaszkiewicz, Genetic Local Search for Multi-Objective Combinatorial Optimization, *European Journal of Operational Research*, 137, pp. 50-71, 2002.

[15] M. Vassileva, A Generalized Scalarizing Problem of Multicriteria Optimization. *Comptes Rendus de l'Academie Bulgare des Sciences*, 58, 5, pp. 537 – 544, 2005.

[16] C. A. Coello, Evolutionary Multi-Objective Optimization: A Historical View of the Field, *IEEE Computational Intelligence Magazine*, February 2006, pp. 29-36, 2006.

[17] C. A. Coello and G. B. Lamont, editors, Applications of Multi-Objective Evolutionary Algorithm, World Scientific, Singapore, 2004, ISBN 981-256-106-4, 2004.

[18] T. Hanne and S. Nickel, "A mutiobjective evolutionary algorithm for scheduling and inspection planning in software development projects", European Journal of Operational Research, vol. 167, no. 3, pp. 663-678, Dec. 2005.

[19] R. M. Hubley, E. Zizler and J. C. Roach, "Evolutionary algorithm for the selection of single nucleonide polymorfizm", BMC Bioinformatics, vol. 4, no. 30, July 2003.

[20] A. Molina-Cristobal, I. A. Grifin, P. J. Fleming and D. H. Owen, "Multiobjective controller structure with genetic algorithms", In Proceedings of the 2005 IFAC World Congres of Automatic Control, Prague, Czech Republic, July, 2005.

[21] S. Majumdar, K. Mitra and G. Sardar, "Kinetic analysis and optimization for the caralync estenfication step of PPT polymerization", Macromolecular Theory and Simulation, vol. 14, pp. 49-59, 2005.