



Introduction of Biogeography-Based Programming as a new algorithm for solving problems



Emadaldin Mohammadi Golafshani*

Department of Civil Engineering, Tehran Science and Research Branch, Islamic Azad University, Tehran, Iran

ARTICLE INFO

Keywords:

Biogeography-Based Programming
Machine learning
Automatic programming

ABSTRACT

Application of evolutionary computation techniques is relatively novel for machine learning. Motivated by different types of evolutionary computation techniques, different types of automatic programming were proposed. Biogeography-Based Optimization (BBO) is a new evolutionary algorithm that is inspired by the science of biogeography and has been shown to be competitive to other population-based algorithms. Inspired by biogeography theory and previous results, in this paper Biogeography-Based Programming (BBP) is proposed as a new type of automatic programming for creating polynomial regression models. In order to show the effectiveness of the proposed BBP, a number of experiments were carried out on a suite set of benchmark functions and the results were also compared with several existing automatic programming algorithms. Furthermore, sensitivity analysis was performed for the parameter settings of the proposed BBP. The results indicate that the proposed model is promising in terms of success rate and accuracy and it performs better than other algorithms investigated in this consideration.

© 2015 Elsevier Inc. All rights reserved.

1. Introduction

The machine learning is a branch of artificial intelligence that deals with the study of systems and can be learned from the data. The machine learning methods, essentially inspired from biological learning, are powerful tools for the design of computer programs that are able to automatically learn with experience [1,2]. They extract knowledge, complex patterns and various discriminators from machine readable data without any need to perform the experimental and numerical tests and make intelligent decisions [1]. The major focus of the machine learning research is on data mining problems, difficult-to-program applications, and software applications customizing to the individual user's preferences [1,3]. Application of evolutionary computation techniques is one of the youngest paradigms inside the machine learning research area. These techniques use iterative progress, such as growth or development in a population. This population is then selected in a guided random search using parallel processing to achieve the desired conclusion. Such processes are often inspired by biological mechanisms of evolution [4]. Having been motivated by evolutionary algorithms, researchers successfully have applied automatic programming algorithms to automatically generate programs or equations among the inputs and outputs. Depending on the type of evolutionary computation techniques used to produce variation in the population, different types of automatic programming models have been subsequently proposed [5–15].

Genetic programming (GP) is known as an extension of genetic algorithm (GA) which the solutions are computer programs rather than fixed length binary strings [16]. GP tackles learning problems by means of searching a computer program space for

* Tel. +98 02164542280.

E-mail address: E.Mohammadi.Golafshani@gmail.com

the program that better respects some given functional specifications. At the most abstract level GP is a systematic, domain-independent method for getting computers to solve problems automatically starting from a high-level statement of what needs to be done. In GP, a population of computer programs is evolved; that is, generation by generation, GP stochastically transforms populations of programs into new, hopefully better, populations of programs. The search is performed using GA [5]. Apart from basic GP with standard crossover (SC) operator, there are various GP based techniques with different crossover operators, such as no same mate (NSM) [6], semantics aware crossover (SAC) [7], context aware crossover (CAC) [8], soft brood selection (SBS) [9] and semantic similarity- based crossover (SSC) [7]. Moreover, there are other different improved versions of genetic programming, for example, Cartesian genetic programming (CGP) [10], gene expression programming (GEP) [11] and linear genetic programming (LGP) [12].

Clone selection programming (CSP) is a new paradigm of evolutionary computation based on the biological immune system concepts. It is an extension of artificial immune system (AIS), which is a systematic, domain independent, and intelligent based method to solve regression problems. A specific operation is implemented by an antibody's affinity and a set of probabilistic parameters [13].

Dynamic ant programming (DAP) is a novel method for automatic programming which is based on ant colony optimization and uses dynamically changing pheromone table. Depending on the value of pheromone, the nodes (terminal and nonterminal) are selected. The selection of nodes with high pheromone rate is more probable. In this method, the search space is dynamically changing and the ants discover good solution using portions of solutions, which are of pheromone value [14].

Artificial bee colony programming (ABCP) is another new approach for automatic programming based on the artificial bee colony (ABC) algorithm. Similar to the relation between GP and GA, ABCP is an adaptation of ABC algorithm which deals with the representation of the problem using more complex structures. Similar to other automatic programming models, this approach allows evolving expressions and constants in the same representation and forming the mathematical functions automatically [15].

Biogeography-Based Optimization (BBO) is a novel type of evolutionary computation techniques proposed by Simon, which is inspired by the geographical distribution and migration of species in an ecosystem [17]. In recent years, BBO has been studied and developed comprehensively and it performs better than other widely used heuristic algorithms like genetic algorithms, ant colony optimization, particle swarm optimization, differential evolution, and simulated annealing for some well-known benchmarks [18–26]. Moreover, BBO successfully has been applied in several practical problems, such as sensor selection problems for aircraft engine health diagnostics [17], groundwater possibility retrieval systems [27], and power flow problems [28] and the results indicated on the strength of BBO. However, biogeographical paradigm parallel to other evolutionary algorithms has not yet been proposed. There has been no attempt to use principles of biogeography to automatically create computer programs. In this paper, Biogeography-Based Programming (BBP), inspired by BBO, is proposed as a novel paradigm combining the program-like representation of solutions to symbolic regression problems with the principles and theories of the biogeographical system. However, BBP is not limited to finding an optimized solution for a specific problem as in BBO algorithm; it is a domain-independent approach in which solutions (computer programs) are generated that can, in turn, solve an entire class of similar problems.

The rest of this study is prepared as follows: In Section 2, the brief overview of BBO algorithm is described. Section 3 presents the proposed BBP. In Section 4, the performance of BBP has been tested on 10 well known benchmark data sets and the results are compared with some existing algorithms. Also, the analysis of performance sensitivity of the parameter settings for the proposed BBP is presented in this section. Finally, concluding remarks are made in Section 5.

2. Biogeography-Based Optimization (BBO)

Biogeography science is defined as the study of the distribution of species and ecosystems over the surface of the earth, in both space and time [24,29–32]. The distribution of species across the surface of the earth usually depends on a combination of environmental reasons. In the natural world, species tend to explore more suitable environments. A good habitat tends to have a large number of species, has a high suitability index (HSI) and vice versa. During the progress of evolution, habitats with a high HSI have a low species immigration rate since they are already nearly saturated with species. Over time, the habitats with high HSI come to have more species, while those with low HSI come to have fewer. Inspired by this idea, Dan Simon proposed a new evolutionary algorithm named Biogeography-Based Optimization (BBO) [17]. Mathematically speaking, a habitat presents a possible solution for the optimization problem. BBO utilizes mathematical models to describe the biogeographical behavior using migration, mutation and the distribution of species [32]. In the following, the main components of BBO are explained.

Migration is a probabilistic operator. The migration rates of each solution are used to modify existing solution by sharing features within the population. For each feature of a given solution y_k , the immigration rate λ_k of y_k determines the probability of immigration of current solution. If the solution y_k is selected for immigration, then the emigrating solution y_j is probabilistically chosen based on the emigration rate μ_j . Migration is written as [33]:

$$y_k(s) \leftarrow y_j(s) \quad (1)$$

where s is a solution feature. Immigration and emigration rates are functions of the number of species in the habitat, such as the linear migration curves in Fig. 1. In addition, some nonlinear migration curves are discussed in Section 5.

Mutation is a probabilistic operator that randomly modifies a solution feature. The purpose of mutation is to enhance the diversity of the population which helps to decrease the chances of getting trapped in local optima [24,33].

Fig. 2 presents the pseudo code of BBO, where $\text{rand}(0,1)$ generates a random real number in the range of [0,1].

Download English Version:

<https://daneshyari.com/en/article/4626226>

Download Persian Version:

<https://daneshyari.com/article/4626226>

[Daneshyari.com](https://daneshyari.com)