# An Eco-inspired Evolutionary Algorithm Applied to Numerical Optimization

Rafael Stubs Parpinelli Applied Cognitive Computing Group Santa Catarina State University Joinville, Brazil Email: parpinelli@joinville.udesc.br

Abstract—The search for nature-inspired ideas, models and computational paradigms always was of great interest for computer scientists, particularly for those from the Natural Computing area. The concept of optimization is present in several natural processes as in the evolution of species, in the behavior of social groups, in the dynamics of the immune system, in the food search strategies and ecological relationships of different animal populations. This work uses the ecological concepts of habitats, ecological relationships and ecological successions to build an ecology-inspired optimization algorithm, named ECO. The proposed approach uses several populations of candidate solutions that cooperates and coevolves with each other, according to a given meta-heuristic. In this particular work, we used the Artificial Bee Colony (ABC) algorithm as the main meta-heuristic. Experiments were done for optimizing benchmarck mathematical functions. Results were compared with the ABC algorithm running without the ecology concepts previously mentioned. The ECO algorithm performed significantly better than the ABC, especially as the dimensionality of the functions increase, possibly thanks to the ecological interactions (intra and inter-habitats) that enabled the coevolution of populations. Results suggest that the eco-inspired algorithm can be an interesting alternative for numerical optimization.

*Keywords*-optimization; cooperative search; co-evolution; habitats; ecological relationships; ecology

#### I. INTRODUCTION

The nature has always been an endless source of inspiration for computational models and paradigms, in particular for the computer scientists of the area known as Natural Computing [1]. The two main families of bio-inspired algorithms are the evolutionary computation (EC) and the swarm intelligence (SI) that currently offer a wide range of strategies for optimization [2], [3].

The concept of optimization is present in several natural processes such as in the evolution of the species, in the behavior of social groups, in the dynamics of the immune system, in the strategies of search for food and in the ecological relationships of different populations. The majority of these cases already serve as inspiration to the development of algorithms for optimization.

However, until now, the ecological concepts of habitats, ecological relationships and ecological succession has not been explored computationally in a context of optimization Heitor Silvério Lopes Bioinformatics Laboratory Federal Technological University of Paraná Curitiba, Brazil Email: hslopes@utfpr.edu.br

[4], [5]. Hence, this work proposes an ecology-inspired algorithm to solve numerical optimization problems.

The main differential of the proposed approach is the cooperative use of different populations of candidate solutions, coevolving in an ecological context. Each population behaves in accordance to a specific search strategy, employed in the evolution of candidate solutions. In addition to the possibility of using different search strategies cooperatively, the methodology opens the possibility for insertion of ecological concepts in the optimization process.

### II. PROPOSED ECOLOGICAL-INSPIRED APPROACH

The ecological-inspired algorithm, named ECO, represents a new perspective to develop cooperative search algorithms. The ECO is composed by populations of individuals (candidate solutions for a problem being solved) and each population evolves according to a search strategy. In this way, the individuals of each population are modified according to the mechanisms of intensification and diversification, and the initial parameters, specific to the search strategy being employed. The eco-inspired system can behave in two ways: homogeneous or heterogeneous. A homogeneous behavior implies that all populations evolve in accordance with the same optimization strategy, configured with the same parameters. Any change in the strategies or parameters in at least one population characterises a heterogeneous behavior.

The ecological inspiration stems from the use of some ecological concepts, such as: habitats, ecological relationships and ecological successions [4], [5]. Once dispersed in the search space, populations of individuals established in the same region constitute an ecological habitat. Thus, a habitat is a set whose elements are populations that belong to the same region of the search space. For example, in a multimodal hyper-surface, each peak can become a promising habitat for some populations. A hyper-surface may possess various habitats. As well as in nature, the populations can move around through all the environment. However, each population may belong only to one habitat in a given moment of time t. Therefore, by definition, the intersection between all habitats at moment t is the empty set.

With the definition of habitats, two categories of ecological relationships can be defined. Intra-habitats relationships that occur between populations inside each habitat and interhabitats relationships that occur between habitats [4], [5].

In ECO, the intra-habitat relationship is the mating between individuals. Populations belonging to the same habitat can establish a reproductive link between their individuals, meshing the populations and favoring the co-evolution of the involved populations. Populations belonging to different habitats are called reproductively isolated.

The inter-habitats relationship is the great migrations. Individuals belonging to a given habitat migrate to other habitats aiming at identifying promising areas for survival and mating.

In addition to the mechanisms of intensification and diversification specific to each search strategy, when considering the ecological context of the proposed algorithm, the intrahabitats relationships are responsible for intensifying the search and the inter-habitats relationships are responsible for diversifying the search.

Inside the ecological metaphor, the ecological successions represent the transformational process of the system. In this process, populational groups are formed (habitats), relations between populations are established and the system stabilizes through the self-organization of its components.

Algorithm 1 shows the pseudo-code of the proposed approach. First of all, at instant t = 0, all individuals of all populations  $Q_i$ , with  $i = 1, \ldots, NQ$ , are randomly initialized (line 2). NQ is the total number of populations.

The main loop represents the ecological successions (lines 3 to 12). The first step inside the main loop is the evolutive period for each population  $Q_i(t)$  (line 4). In this step each population explores the search area freely in accordance to its intensification and diversification criteria.

At the end of the evolutive period of all populations it is necessary to identify the region of reference for each population (line 5). The region of reference is used to define the habitats of the system. The metric chosen to define the region of reference is the centroid  $C_i$ , with i = 1, ..., NQ. The centroid represents the point in the space where there is a longest concentration of individuals of population i. Once found the centroids, the distances between populations are calculated and the habitats  $H_j(t)$  are defined using a minimum distance threshold  $\rho \in [0..1]$  (line 6). The distance between populations is dependent on the definition and representation of each population. In this application we used the Euclidean distance. At the end of this step all populations are associated with a habitat:  $Q_i^j(t)$ , population i belongs to the habitat j at the moment t.

With the *NH* habitats defined, it is necessary to determine the intra-habitats communication topologies  $CT_j(t)$  for each habitat  $H_j(t)$  with j = 1, ..., NH (line 7). The topologies

## Algorithm 1 Pseudo-code for ECO

- 1: Consider i = 1, ..., NQ, j = 1, ..., NH and t = 0;
- 2: Initialize each population  $Q_i(t)$  with  $n_i$  random candidate solutions;
- 3: while stop criteria not satisfied do {Ecological succession cycles}
- 4: Perform evolutive period for each population  $Q_i(t)$ ;
- 5: Apply metric  $C_i$  to identify the region of reference for each population  $Q_i(t)$ ;
- 6: Using the  $C_i$  values, define the NH habitats;
- 7: For each habitat  $H_j(t)$  define the communication topology  $CT_j(t)$  between populations  $Q_i^j(t)$ ;
- 8: For each topology  $CT_j(t)$ , perform interactions between populations  $Q_i^j(t)$ ;
- 9: Define communication topology TH(t) between  $H_i(t)$  habitats;
- 10: For TH(t) topology, perform interactions between  $H_i(t)$  habitats;
- 11: Increase t;
- 12: end while

are defined by an adjacency matrix between the populations belonging to each habitat. Once defined the topologies intrahabitats, the mating ecologic relationship between adjacent populations occurs (line 8). In this relationship, one individual of each population is chosen using the tournament strategy [6] and a genetic exchange between them is performed in order to generate a new individual. The new generated individual replaces an individual selected at random in its population, excluding the best individual.

Once the interactions between the populations of each habitat have been done, the TH(t) topology for interaction among habitats (line 9) is randomly defined. This interhabitats topology TH(t) is used for the completion of the great migrations ecologic relationship. In this relationship, for each habitat a random population belonging to him is chosen. The best individual of this population migrates to another habitat and, in the destination habitat, it replaces an individual randomly chosen, excluding the best individual (line 10).

The main loop continues until the ecological succession cycles reach the maximum predefined value.

#### **III. EXPERIMENTS AND RESULTS**

Experiments were conducted using four benchmark functions extensively used in the literature for testing optimization methods [7]. Each function was tested with 2, 5, 10, and 200 dimensions and, for all of them, the objective is to minimize.

The first function  $(f_1(\vec{x}))$  is known as generalized F6 Schaffer function [8] that is strongly multimodal. The second function  $(f_2(\vec{x}))$  is the Rastrigin function and it is also multimodal and based on the sphere function with the addition of a cosinoid modulation to produce a large number of local minima [9]. The third function  $(f_3(\vec{x}))$  is the Griewank function [10] that, as the other, is strongly multimodal. Particularly the number of local minima for this function increases exponentially with the dimensionality [11], which makes it a very difficult optimization problem. The fourth function  $(f_4(\vec{x}))$  is the Rosenbrock function [12], which has the particularity of having a search hyper-surface in the form of a long and narrow parabolic valley, where the global optimum is located. Table I summarizes the informations about the functions used.

The parameters of the ECO algorithm are: number of populations (*N-POP*) that will be co-evolved, number of cycles for ecological successions (*ECO-STEP*), the size of the evolutive period (*EVO-STEP*), the tournament size (*T-SIZE*) and the proximity threshold  $\rho$ . Studies about the adjustment of parameters have not been carried out yet. Hence, all the parameters of the algorithm were defined empirically.

In all experiments the Artificial Bee Colony Optimization (ABC) algorithm [13] was used in a homogeneous manner, i.e. all populations use this algorithm with the same adjustment of parameters to evolve their candidate solutions. For the ABC algorithm, there is only one control parameter, the population size (*POP-SIZE*).

#### A. Conceptual Illustration

The generalized Schaffer function  $(f_1(\vec{x}))$  was chosen to illustrate in details the behavior of the proposed algorithm. With the purpose of allowing a visual assessment of the behavior of the algorithm, the Schaffer function was defined here with only two dimensions. In this illustration, the parameters used were *N*-*POP* = 10, *POP*-*SIZE* = 10, *ECO*-*STEP* = 100, *EVO*-*STEP* = 5, *T*-*SIZE* = 5 e  $\rho$  = 0, 5. With this configuration, the total number of function evaluations is 10,000 for each population. A total of 100 evaluations for each population was done in each ecological succession. The parameters were chosen empirically.

First of all, all populations are randomly initiated. The initialization uses a normal distribution  $(N(\mu, \sigma))$  with average  $\mu$  (randomly chosen within the domain of each dimension of the problem) and standard deviation  $\sigma$  (also proportional to the problem domain). The random sampling is done using the Box-Muller transformed [14].

Figure 1(a) shows the initial distribution of the individuals of all *N-POP* populations: *Sp*0, ..., *Sp*9.

Once initialized the populations, the algorithm enters the ecological succession loop that begins with the evolutive period. In this stage, all populations evolve their solutions for a pre-established number of iterations. In this example, the evolutive period was defined as 5 iterations.

After the evolutive period, begins the definition of habitats. The habitats are the regions in which the populations are concentrated. In this proposal, the region of reference of a population is defined by its centroid and is calculated by Equation 1, where *Pop* represents the number of individuals in the population and  $\vec{x}_i$  represents an individual (solution vector). Figure 1(b) illustrates the distribution of all centroids of all populations for the first ecological succession.

$$C = \frac{\sum_{i=1}^{P_{op}} \vec{x}_i}{P_{op}} \tag{1}$$

Once found the centroids for each population, the Euclidean distance between them is calculated and the habitats are defined in accordance with the minimum threshold  $\rho$ . The adjacency matrix is generated according to  $\rho$ . Two populations are called adjacent among themselves if they are at a distance of at least  $\rho$  one of another. In Figure 1(b) all information of adjacency between centroids are shown in arrows: Sp0 is not adjacent to Sp7; Sp2 is adjacent to Sp7; Sp3 is adjacent to Sp6; Sp4 is not adjacent to any population; Sp1 is adjacent to Sp1 and Sp2; Sp8 is not adjacent to any population; and Sp9 is not adjacent to any population.

The habitats are generated using the adjacency matrix. Figure 1(c) illustrates the habitats found for the centroids distribution of Figure 1(b), with  $\rho = 0.5$ . It is observed in this figure the existence of seven habitats:  $H_0$  composed by Sp0;  $H_1$  composed by Sp1, Sp2 and Sp7;  $H_2$  composed by Sp3 and Sp6;  $H_3$  composed by Sp4;  $H_4$  composed by Sp5;  $H_5$  composed by Sp8; and  $H_6$  composed by Sp9.

The adjacency matrix also defines the intra-habitat communication topology. For example, in habitat  $H_1$ , composed by populations Sp1, Sp2 and Sp7, the population Sp1 can establish a relationship with the population Sp7, the population Sp2 can establish a relationship with the population Sp7, and the population Sp7 can establish a relationship with populations Sp1 and Sp2. At this moment the habitats are well defined with their populations and communication topologies.

The next step is the communication between populations within each habitat. Populations that are adjacent between themselves interact by the mating ecologic relationship. Consider the habitat  $H_1$ . The population Sp7 selects an individual of its population and an adjacent population to establish a relationship. The adjacent population is chosen at random and, in this case, could be the population Sp1 or Sp2. In each population, the individuals chosen to carry out mating are selected by using the tournament selection of size 5. The new generated individual replaces an individual randomly chosen within the adjacent population. The populations Sp1 and Sp2 perform the same procedure. All habitats composed by more than one population carry out mating according to the topology defined by the adjacency matrix.

After the intra-habitats interactions it is necessary to define the inter-habitats communication topology. This topol-



Table I BENCHMARK FUNCTIONS

Figure 1. Conceptual illustration of the proposed algorithm.

ogy is used to perform the great migrations ecologic relationship. In this relationship, for each habitat a random population belonging to him is chosen at random. The best individual of the population chosen migrates to another random habitat and, in the destination habitat, it replaces an individual chosen at random, excluding the best individual. In this stage of the algorithm the ecological succession loop restarts. Figure 1(d) illustrates the habitats found in the last ecological succession step. They are:  $H_0$  composed by Sp0, Sp1, Sp2, Sp3, Sp4, Sp5, Sp7 and Sp9;  $H_1$  composed by Sp6; and  $H_2$  composed by Sp8. We can also observe in Figure 1(d) that populations belonging to the habitat  $H_0$  converged toward the global optimum and other habitats have converged to regions close to the global optimum. Figure 2 shows the evolution of the number of habitats for each ecological succession step. It is observed that the system has evolved and converged to the formation of three habitats.

#### B. Numerical Optimization of Benchmark Functions

The eco-inspired algorithm (ECO) was tested using two configurations. One configuration implements the Algo-



Figure 2. Number of habitats at each ecological succession step.

rithm 1 as described in Section II, with the definitions of habitats, topologies and ecological relations. In the other configuration the ability to create habitats is disabled and, consequently, topologies and interactions are not defined. This second configuration simulates the evolution of the populations completely isolated. For each configuration, the algorithm was run 30 times.

For the number of dimensions equal to 2, 5 and 10, there was only one parameter change related to the previous illustration. For these experiments, the number of populations was defined as 100. For the number of dimensions equal to 200, given the high number of dimensions, the number of populations was redefined to 200, the number of ecological succession cycles was redefined to 500, and the evolutionary period was redefined to 10 iterations. With this adjustment of parameters, for 200 dimensions, the total number of function evaluations to each configuration of the algorithm has become from 10,000 to 100,000 evaluations for each population were done at each ecological succession. The parameters were chosen empirically.

Table II shows the averaged results obtained for the benchmark functions. For each function, the first line shows the number of dimensions: 2, 5, 10 and 200 respectively. The second line presents the configuration of the algorithm, with (ECO) and without (ECO-no) the use of ecological concepts, respectively. The ECO-no usage is equivalent to run the ABC algorithm without co-evolution, where the populations evolve without exchange of information. The third line (Avg best) shows the average and standard deviation of the best results for all populations in each run. In this experiment, each run of the algorithm gives the evolution of all populations, each with its convergence. At the end of a single run we have the average of the best results of all populations. At the end of 30 executions it is calculated the average of averages. The fourth line (Global Best) shows the average and standard deviation of the best result obtained by all populations in all runs.



Figure 3. Average number of habitats at each ecological succession.

Analyzing Table II we can observe that the proposed algorithm ECO obtained significantly better results than the algorithm being executed without the concepts of habitat and interactions between populations (ECO-no), for all functions. This gain can be best observed with the increase of dimensionality of each function and it is mainly due to the ecological interactions (intra and inter-habitats) that favor the co-evolution of populations.

Figure 3 shows the evolution of the average number of habitats for each ecological succession step for 30 executions of the ECO algorithm. As the observed behavior is the same for all functions, Figure 3 shows the evolution for the function  $f_1(\vec{x})$  with D = 10. It can be observed that, at the beginning of the optimization process, with the populations widely dispersed in the search space, there is a greater number of habitats. To the extent that the optimization process moves through the ecological successions, the populations tend to move through the search space converging to specific regions. As shown in Figure 3, the number of habitats decreases with the ecological succession cycles, indicating that the populations tend to converge to points close to each other.

#### **IV. CONCLUSIONS**

This paper presents an eco-inspired algorithm for optimization, called ECO. The proposed algorithm uses cooperative search strategies where populations of individuals coevolve and interact among themselves using some ecological concepts. Each population behaves according to the mechanisms of intensification and diversification, and the control parameters, specific to a search strategy. In this work was used the Artificial Bee Colony Optimization algorithm in all populations.

The main ecological concepts addressed are the definition of habitats, ecological relationships and ecological successions. Along with these concepts, the definition of intra and inter-habitats communication topologies are inserted to compose the algorithm. These features bring a greater

 Table II

 Obtained results for the benchmarck functions.

$f_1(\vec{x})$	D=2		D = 5		D = 10		D = 200	
Model	ECO	ECO-no	ECO	ECO-no	ECO	ECO-no	ECO	ECO-no
Avg Best	$0.0118 \pm 0.0$	$1.7345 \pm 0.6$	$0.5239 \pm 0.1$	$9.6695 \pm 0.4$	$3.9895 \pm 0.3$	$12.6312 \pm 0.3$	$20.5058 \pm 0.2$	$28.5370 \pm 0.8$
Global Best	$0.0000 \pm 0.0$	$0.0186 \pm 0.1$	$0.1341 \pm 0.0$	$3.7867 \pm 1.5$	$2.6070 \pm 0.2$	$10.0630 \pm 0.9$	$18.3801 \pm 0.4$	$27.5936 \pm 0.7$
$f_2(\vec{x})$	D=2		D = 5		D = 10		D = 200	
Model	ECO	ECO-no	ECO	ECO-no	ECO	ECO-no	ECO	ECO-no
Avg Best	$0.0022 \pm 0.0$	$0.8426 \pm 0.2$	$0.1694 \pm 0.2$	$12.2897 \pm 3.8$	$1.2263 \pm 1.0$	$55.8775 \pm 3.5$	$4.6851 \pm 1.7$	$208.6492 \pm 2.2$
Global Best	$0.0000 \pm 0.0$	$0.0000\pm0.0$	$0.0000 \pm 0.0$	$0.3832 \pm 0.6$	$0.0000 \pm 0.0$	$12.6207 \pm 4.3$	$0.0003 \pm 0.0$	$136.3832 \pm 6.4$
$f_3(\vec{x})$	D = 2		D = 5		D = 10		D = 200	
Model	ECO	ECO-no	ECO	ECO-no	ECO	ECO-no	ECO	ECO-no
Avg Best	$10^{-17} \pm 0.0$	$10^{-17} \pm 0.0$	$0.0012 \pm 0.0$	$0.0016 \pm 0.0$	$0.0002 \pm 0.0$	$1.1613 \pm 1.2$	$10^{-10} \pm 0.0$	$10^{-7} \pm 0.0$
Global Best	$0.0000\pm0.0$	$10^{-19} \pm 0.0$	$10^{-19} \pm 0.0$	$10^{-12} \pm 0.0$	$10^{-13} \pm 0.0$	$10^{-06} \pm 0.0$	$10^{-11} \pm 0.0$	$10^{-8} \pm 0.0$
$f_4(\vec{x})$	D = 2		D = 5		D = 10		D = 200	
Model	ECO	ECO-no	ECO	ECO-no	ECO	ECO-no	ECO	ECO-no
Avg Best	$0.0000 \pm 0.0$	$4.9315 \pm 0.6$	$1.0332 \pm 1.5$	$369.5634 \pm 219.3$	$87141.5 \pm 1562.8$	$2.65 \times 10^6 \pm 3421.2$	$333.0645 \pm 120.0$	$5.19 \times 10^6 \pm 21664.5$
Global Best	$0.0000 \pm 0.0$	$0.0015\pm0.0$	$0.0019\pm0.0$	$0.4379 \pm 0.2$	$0.0646 \pm 0.0$	$93.2517\pm62.7$	$137.8643 \pm 42.0$	$13036.1\pm4193.4$

biological plausibility to the proposed algorithm.

Experiments were carried out considering the implementation of the algorithm with and without the use of ecological concepts with different benchmark functions, each with different degrees of complexity. The results showed that the use of habitats and ecological relations influence significantly the co-evolution process of populations, leading to better solutions (when compared to the results not using the ecological concepts).

This work is still in development and as future work we intend to analyze the influence of the system parameters (number of ecological successions, evolutive period, number of populations, and  $\rho$  threshold for creation of habitats) on the quality of solutions, as well as to use different search strategies in the proposed model (heterogeneous system). This last research branch aims at analyzing the effect that each search strategy can have on the results, keeping in mind that each one has its own way to explore the solutions space. Currently, in order to increase the computational capabilities of the proposed algorithm, a parallel framework is under development.

#### ACKNOWLEDGMENT

Authors would like to thank the Brazilian National Research Council (CNPq) for the research to H.S. Lopes; as well as to UDESC (Santa Catarina State University) and FUMDES program for the doctoral scholarship to R.S. Parpinelli.

#### REFERENCES

- [1] L. N. de Castro, "Fundamentals of natural computing: an overview," *Physics of Life Reviews*, vol. 4, no. 1, pp. 1–36, 2007.
- [2] A. P. Engelbrecht, *Computational Intelligence: An Introduction*, 2nd ed. Chichester, UK: Wiley, 2007.
- [3] R. S. Parpinelli and H. S. Lopes, "New inspirations in swarm intelligence: a survey," *International Journal of Bio-Inspired Computation*, vol. 3, no. 1, pp. 1–16, 2011.

- [4] M. Begon, C. R. Townsend, and J. L. Harper, *Ecology: from individuals to ecosystems*, 4th ed. Oxford, UK: Blackwell Publishing, 2006.
- [5] R. M. C. May and A. R. McLean, *Theoretical Ecology: Principles and Applications*. Oxford, UK: Oxford University Press, 2007.
- [6] T. Blickle, "Tournament selection," in *Evolutionary Computation*, T. Bäck, D. Fogel, and Z. Michalewicz, Eds. Bristol, UK: Institute of Physics, 2000, vol. 2, pp. 181–186.
- [7] J. G. Digalakis and K. G. Margaritis, "An experimental study of benchmarking functions for evolutionary algorithms," *International Journal of Computer Mathematics*, vol. 79, no. 4, pp. 403–416, 2002.
- [8] C. A. Floudas and P. M. Pardalos, A collection of test problems for constrained global optimization problems, ser. Lecture Notes in Computer Science. Springer, 1990, vol. 455.
- [9] H. Mühlenbein, D. Schomisch, and J. Born, "The parallel genetic algorithm as function optimizer," *Parallel Computing*, vol. 17, no. 6-7, pp. 619–632, 1991.
- [10] A. Griewank, "Generalized descent for global optimization," *Journal of Optimization Theory and Applications*, vol. 34, no. 1, pp. 11–39, 1981.
- [11] H. Cho, F. Olivera, and S. Guikema, "A derivation of the number of minima of the Griewank function," *Applied Mathematics and Computation*, vol. 204, no. 2, pp. 694–701, 2008.
- [12] H. Rosenbrock, "An automatic method for finding the greatest or least value of a function," *The Computer Journal*, vol. 3, pp. 175–184, 1960.
- [13] D. Karaboga and B. Akay, "A comparative study of artificial bee colony algorithm," *Applied Mathematics and Computation*, vol. 214, pp. 108–132, 2009.
- [14] G. E. P. Box and M. E. Muller, "A note on the generation of random normal deviates," *The Annals of Mathematical Statistics*, vol. 29, no. 2, pp. 610–611, 1958.