The Baldwin Effect on a Memetic Differential Evolution for Constrained Numerical Optimization Problems

Saul Dominguez-Isidro and Efren Mezura-Montes
Artificial Intelligence Research Center
University of Veracruz
Sebastian Camacho No. 5
Xalapa, Veracruz, Mexico 91000
sdominguezisidro@gmail.com,emezura@uv.mx

ABSTRACT

This paper analyzes the Baldwin effect on a memetic algorithm that solves constrained numerical optimization problems (CNOPS). For this study, the canonical Differential Evolution (DE) enhanced with the Hooke-Jeeves method (HJ) as local search operator is proposed (MDEHJ), which implements a probabilistic scheme to activate HJ by means a sinusoidal function that considers the population diversity. Three MDEHJ instances are applied to study the Baldwin effect in different exploitation areas (best, worst and random selected, respectively). Final results are compared against those obtained by MDEHJ with Lamarckian learning. All instances are tested on thirty-six well-known benchmark problems. The results suggest that the proposed approach is suitable to solve CNOPS and those results also show that Baldwin effect does not affect the performance of a memetic DE in constrained search spaces.

KEYWORDS

Memetic Algorithm, Differential Evolution, Constrained Numerical Optimization, Baldwin Effect

ACM Reference format:
Saul Dominguez-Isidro and Efren Mezura-Montes. 2017. The Baldwin Effect on a Memetic Differential Evolution for Constrained Numerical Optimization Problems. In *Proceedings of GECCO '17 Companion, Berlin, Germany, July 15-19, 2017,* 2 pages.
DOI: http://dx.doi.org/10.1145/3067695.3076096

1 INTRODUCTION

Memetic algorithms (MAs) incorporate a set of local improvement procedures into an evolutionary algorithm, to help global search find better results by exploiting promising regions of the search space through what could be termed "learning". Two elemental models of evolution can be used to incorporate learning into a MA: the Lamarckian learning and Baldwin Effect. While Baldwin effect, just modifies the fitness value (phenotype) of a solution, i.e., the result of the improvement does not change the structure of the solution (genotype); the Lamarckian learning transforms the phenotype and genotype of a solution. According to previous studies, the Baldwin effect can has benefits in flat landscapes and can avoid premature convergence in hard problems. However, the Baldwin effect in MAs for CNOPs is not usually analyzed to the best of the authors' knowledge. Therefore, this is the main motivation for this work which is aimed to analyze the Baldwin effect into an MA based on Differential Evolution (DE) [4] that incorporates the Hooke-Jeeves (HJ) [2] method as local search operator. The ε-constrained method is adopted as constraint-handling [5].

GECCO '17 Companion, Berlin, Germany

DOI: http://dx.doi.org/10.1145/3067695.3076096

2 PROPOSED APPROACH

In this proposal, HJ is handling by means a probability mechanism, which includes the population diversity information regarding the objective function values. The probability is computed as follows:

$$P = p_{min} * (sin(2\pi * freq * G) * \chi + 1) \tag{1}$$

where p_{min} is the minimum probability allowed to be applied the local search, $freq$ is the frequency of probability variation, while G is the current generation of DE framework. Finally, χ (see Eq. 2) is an estimation of the best vector performance with respect to the other vectors [1].

$$\chi = \frac{|f_{best} - f_{avg}|}{max|f_{best} - f_{avg}|_G} \tag{2}$$

where f_{best} and f_{avg} are the objective function values of the best and average solutions of the population, respectively. $max|f_{best} - f_{avg}|_G$ is the maximum difference observed (e.g., at the generation G). According to Equation 1, if the population tends to resemble the best performance vector, the probability P to activate the local searcher tends to p_{min}. Algorithm 1 denotes the whole procedure, where ϕ represents the constraint violation sum of a solution.

3 EXPERIMENTS AND RESULTS

Two experiments were designed to analyze the Baldwin effect on CNOPs by using a memetic DE, and consist of analyzing three MDEHJ instances with different exploitation areas (best, worst and random selected, respectively). All cases were tested on eighteen well-known benchmark problems [3] with two search space dimensionality (10D and 30D). Statistical values of 25 independent runs per test function were performed and compared against those results obtained by MDEHJ instances that use Lamarckian learning. The following parameter values were used by the components of MDEHJ: For DE, maximum number of population $Pmax$ = 80, crossover probability Cr = 0.9, scaling factor F = 0.55, and maximum number of fitness evaluation $MaxFEs$ = $2.0E + 05$ and $6.0E + 05$ in 10D and 30D, respectively. For HJ, reduction factor α =rand(2, 3), step size δ = rand(0.75, 0.9), and $MaxFEs$ = 700. For ε-Constrained, θ = 0.75, cp = 9.5, and Gc = 1100. For activation mechanism, p_{min} =

Algorithm 1 Memetic Differential Evolution MDEHJ

Randomly generate an initial population of vectors $P_0 = (X_{0,i}, \ldots, X_{0,Pmax})$
Calculate the fitness of each vector in the initial population.
repeat
 for $i \leftarrow 1, Pmax$ **do**
  Randomly select $r0$,$r1$,$r2 \in [1, Pmax]$ and $r0 \neq r1 \neq r2 \neq i$
  Randomly select $J_{rand} \in [1, D]$
  for $j \leftarrow 1, D$ **do**
   if $rand_j \leq Cr$ Or $j = Jrand$ **then**
    $u_{G,i,j} = x_{G,r0,j} + F(x_{G,r1,j} - x_{G,r2,j})$
   else
    $u_{G,i,j} = x_{G,i,j}$
   end if
  end for
  if $U_{G,i}$ is better than $X_{G,i}$ **then**
   $X_{G+1,i} = U_{G,i}$
  else
   $X_{G+1,i} = X_{G,i}$
  end if
 end for
 Set the population diversity χ using Equation 2
 Set the activation probability P using Equation 1
 if $rand(0, 1) \leq P$ **then**
  Set j^*
  Set $X_{new} \leftarrow$ **Hooke-Jeeves**$(X_{G,j})$
  Set $f(X_{G,j}) \leftarrow f(X_{new})$ and $\phi(X_{G,j}) \leftarrow \phi(X_{new})$ (Baldwin Effect)
 end if
 $G = G + 1$
until *MaxFEs* is reached

0.05. The first experiment analyzes the performance between the MDEHJ instances with Baldwin effect and Lamarckian learning. 95%-confidence Wilcoxon rank-sum-test was computed to determine the statistically significant difference between approaches. The comparison is summarized in Table 1. The approximation sign ("≈") determines that there is no significant difference between the two algorithms. The results show that the Baldwin effect does not

Table 1: 95%-confidence Wilcoxon rank-sum-test for the MDEHJ instances with Baldwin effect ($B\text{-}MDEHJ_{best}$, $B\text{-}MDEHJ_{worst}$ and $B\text{-}MDEHJ_{rand}$) against MDEHJ instances with Lamarckian learning ($L\text{-}MDEHJ_{best}$, $L\text{-}MDEHJ_{worst}$ and $L\text{-}MDEHJ_{rand}$). Dim means the dimension of results, while $w+$ and $w-$ mean the sum of positive and negative ranks, respectively. Finally, Diff denotes whether there is a significant difference.

Algorithms	Dim	Criteria	$w+$	$w-$	Diff
$B\text{-}MDEHJ_{best}$ to $L\text{-}MDEHJ_{best}$	10	Best	11	17	≈
		Mean	89	64	≈
	30	Best	103	68	≈
		Mean	93	78	≈
$B\text{-}MDEHJ_{worst}$ to $L\text{-}MDEHJ_{worst}$	10	Best	18	18	≈
		Mean	71	65	≈
	30	Best	110	43	≈
		Mean	107	46	≈
$B\text{-}MDEHJ_{rand}$ to $L\text{-}MDEHJ_{rand}$	10	Best	16	5	≈
		Mean	84	52	≈
	30	Best	116	55	≈
		Mean	126	45	≈

affect the final performance of the algorithm. Although there is no statistical difference between the MDEHJ instances, the numerical results suggest that Baldwin effect promotes the search for those functions where the feasible space is small. The second experiment studies the behavior of Baldwin effect regarding the feasibility rate from the proposed MDEHJ instances ($MDEHJ_{best}$, $MDEHJ_{worst}$, and $MDEHJ_{rand}$). Of thirty-six problems (10D and 30D), MDEHJ samples were able to find 100% feasibility of the 25 runs in most cases. Only C16 function in 10D cases of MDEHJ using the Baldwin effect generate between 13% and 16% feasible solutions. In contrast, the instance with Lamarckian learning that exploits the best solution of the population did not perform well in problems C02

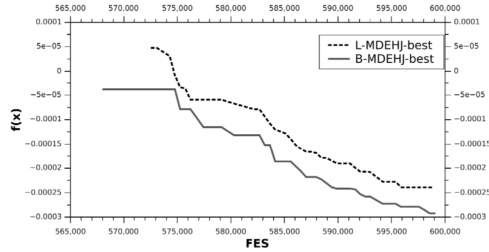

Figure 1: Convergence plot of the median run for C11 test problem, 30D.

and C16 in 10D generating only one feasible run, whereas in 30D problems C09 and C10 were not able to find feasible solutions. The results suggest that the Baldwin effect has a greater influence on the process by exploiting the best area of the population since it can transform the fitness function and constraint violation sum landscape into flat landscapes nearby the local optimal. Despite this feature leads to a slower convergence in most problems than using Lamarckian learning, through exploiting the best solution, the MDEHJ that applies the Baldwin effect can outperform the convergence velocity, see Figure 1.

4 CONCLUSIONS

This work presented an analysis of Baldwin Effect on a memetic differential evolution to solve constrained numerical optimization problems. The results of experiments suggested that the proposed approach was suitable to solve CNOPS and those results also showed that Baldwin effect did not affect the performance of a memetic DE in constrained search spaces. Despite there was not statistically difference considering all test problems, the behavior of memetic approaches depends on the exploitation area. Whereas promising areas were favored by the Baldwin effect, the randomly chosen zones and worst regions were adequately utilized applying Lamarckian learning, since it disrupt the evolutionary cycle modifying the genotype and phenotype information, causing a rapid convergence. The future work consists on hybridizing both learning mechanisms to obtain its advantages.

ACKNOWLEDGMENTS

The authors acknowledge support from CONACYT project No. 220522.

REFERENCES

[1] Andrea Caponio, Ferrante Neri, and Ville Tirronen. 2009. Super-fit control adaptation in memetic differential evolution frameworks. *Soft Computing* 13, 8-9 (2009), 811–831. DOI: http://dx.doi.org/10.1007/s00500-008-0357-1

[2] Robert Hooke and T. A. Jeeves. 1961. "Direct Search" Solution of Numerical and Statistical Problems. *J. ACM* 8, 2 (1961), 212–229. DOI: http://dx.doi.org/10.1145/321062.321069

[3] R. Mallipeddi and P.N. Suganthan. 2010. *Problem Definitions and Evaluation Criteria for the CEC 2010 Competition on Constrained Real-Parameter Optimization.* Technical Report. Nanyang Technological University, Singapore.

[4] Rainer Storn and K Price. 1997. Differential evolution a simple and efficient heuristic for global optimization over continuous spaces. *Journal of global optimization* (1997), 341–359. DOI: http://dx.doi.org/10.1023/A:1008202821328

[5] Tetsuyuki Takahama, Setsuko Sakai, and Noriyuki Iwane. 2005. Constrained Optimization by the Constrained Hybrid Algorithm of Particle Swarm Optimization and Genetic Algorithm. In *AI 2005: Advances in Artificial Intelligence*, Shichao Zhang and Ray Jarvis (Eds.). Lecture Notes in Computer Science, Vol. 3809. Springer Berlin Heidelberg, 389–400.