*International Journal of Advances in Computer Science & Its Applications – IJCSIA*
*Volume 4: Issue 4 [ISSN 2250-3765]*
*Publication Date : 27 December,2014*

# Non-Standard Approaches of Evolutionary Algorithms in the Exploration-Exploitation Dilemma

Oswaldo E. Vélez-Langs

***Abstract--* This critical-reflexive paper is based on the known problem presented in the Exploration-Exploitation Dilemma (EED) when working with Evolutionary Algorithms (EAs) and focusing on proposals that put forward, in this case, both a traditional approach as well as more recent approaches that manage population solutions (individuals) differently than the standard EAs, that is: The Learnable Evolutionary Model (LEM) and the Distribution Estimation Algorithms (DEA).**

***Keywords-* Exploration-Exploitation Dilemma, Evolutionary Algorithms, Learnable Evolution Model, Estimation of Distribution Algorithms**

## I. Introduction

In an optimization problem, when can it be known that what has been obtained is sufficiently good versus what is yet to be obtained? Undoubtedly there are different models, from the scientific model, regarding how to respond to this dilemma. There is a telling example drawn from the design of statistics experiments that illustrates EED in computer studies ([7]): "There is a slot machine with different levels and a better that plays on that machine; in order to obtain the highest winnings from the slot machine, the better should establish a balance between being able to exploit the level so as to get the highest earnings hoped for, and know how to explore so as to get more information about the winnings hoped for in other levels." This is the dilemma that this essay tries to address from the point of view of EAs and their models; not traditional ones but rather recent models.

Evolutionary Algorithms are meta-heuristic models of optimization and searches based on populations of individuals and are inspired by natural evolution ([23], [4], [10], [11]). AEs are well known for being able to perform effective searches against noisy backgrounds and, similar to the previously illustrated example, all search algorithms need to establish a balance between two apparently opposite factors:

- exploration of space solutions in order to perform a wider search, thus finding promising areas.
- exploitation search space in order to perform a deep search in those areas, thus obtaining the best solutions.

In EAs, exploitation is seen in the so-called "selective pressure" which focuses on promising regions of search space, while exploration tries to maintain diversity, that is, avoid a rapid convergence towards areas that do not contain optimums or global optimums. Diversity is associated with the difference between individuals in a population, since if there is a lack of genetic diversity, all the individuals of a population would be very similar and the lack of diversity entails a premature search of local optimums. Later on we will see that there are certain mechanisms that can be used to alleviate this fact. There is a group of predefined operators in EAs that focus on exploitation, the known operators of selection and breeding. Breeding operators make the interchange of information among individuals easier, whereas there are other exploration operators, and these are mutation operators, that introduce diversity.

From the very point of view of conception and design of the EA, an inappropriate balance between Exploration-Exploitation leads to an inefficient search and therefore to solutions that are not desired or are not useful. In trying to maintain a complete balance in the EED, it makes sense that the operators in charge of selection and breeding appropriately perform what they should do, that is, that each new individual has interesting characteristics from their ancestors (or parents) and that the operator of the mutation allows the algorithm to reach whatever point within the search space, all of the above focused on keeping an adequate diversity, a war horse in the essence of the EAs. Part of the foregoing is exposed within the different approaches, presented below, that have been used to deal with EED, to finally go on to expound some conclusions and open questions

Engineering Department
Universidad de Bogotá – Jorge Tadeo Lozano

respecting the way traditional or non-traditional EED from the EAs are focused.

# II. Development

There is the clear intention that an EA is capable, as far as possible, of refining the Exploration-Exploitation relation, and given that both aspects are related to the use of certain operators whether in breeding or mutation, a clear tendency exists to perform this refining within said parameters. An interesting study that contemplates different approximations to this refinement from auto-adaption is found in ([18]).

Although there can be an attempt to intuitively maintain the parameter associated with mutation, called mutation rate, high, to counteract premature convergence, this is not the most appropriate way, since the algorithm could begin to diverge too much. A more adequate solution to this is adapting the mutation rate and it should be considerable in the beginning, and this should be reduced as the algorithm begins to show convergence; a similar focus is that of applying a high mutation rate to bad solutions and low mutation rates to good solutions.

In relation to breeding, it can also be expected, as was mentioned, to maintain an effective diversity each time said operator is applied. There are at least 5 strategies related to this object: the technique of pairing (parents can be selected in order to maintain the diversity of the population), the technique to generate offspring (here there is an important documentary archive from the GAs with real codification, ([13], [9] and [6]), the number of parents (two or more parents are used to mate), the number of children (two or more children are obtained from the parents).

In general, the variation (adaption) of the breeding and mutation operators can be found in two ways: through a deterministic rule or dynamically. Concrete proposals related to a deterministic variation of breeding and mutation operators can be found in: ([14], [5], [3], [17], [20]). On the other hand there are related proposals with a dynamic and adaptive variation of the operators, such as the 1/5 rule of ([28]), among others we have: ([8]), [36]), [30], [34]).

One of the first approaches of non-standard GA which clearly illustrates the above-mentioned is found in the CHC (cross generational elitist selection heterogeneous recombination cataclysmic mutation algorithm). CHC ([12]) is a non-traditional Genetic Algorithm that uses a very conservative selection strategy: always choose the best individuals to form part of the new population. A highly explorative recombination operator is also applied (HUX) that produces descendants as different as possible from both parents; CHC incorporates a mechanism to reset the algorithm when faced with a premature convergence condition. It also introduces direction to avoid the mating of similar individuals. When the CHC cannot insert descendants in a population in a successive generation and the mating threshold has already reached the value of 0, the CHC injects new diversity into the population be means of a reset known as cataclysm mutation. Cataclysm mutation uses the best individual in a population as a model to reseed the population. The new population includes a copy of the chain model and the rest of the population is generated by mutating some percentage of bits (for example, 35%) in the chain model.

The BGA (Breeder Genetic Algorithm), proposed by ([25]), is a method based on animal breeding. This animal breeding has advanced from an art based on intuition to an empirical science based on statistics. If a human breeder does not have information about the genetic material, he should estimate the added value, the breeding value of an animal. The virtual breeder of BGA knows all the genes of his population and additionally controls genetic operators (mutation and breeding rates). The search process in BGA is principally guided by breeding, transforming it into a Genetic Algorithm. The mutation is a background operator where the reason for mutation is inversely proportional to the number of parameters to be optimized, and the range of mutation is fixed. The BGA tries to respond to some questions related to what can happen in case of a mutation model, in relation to successful mutations, and also what can happen with successful breeding, given a selection and breeding model. Normally mutation and breeding are considered disruptive operators; the BGA considers mutation and breeding as construction operators and evaluates them according to the probability they have of creating better solutions.

[2] present a Cellular Genetic Algorithm (CGA) where individuals are placed on a bi-dimensional grid where it is specified that each individual can only mate with their respective communities and an appropriate mechanism for population diversity is provided, varying the way to specify the community within that bi-dimensional grid, and therefore analyzing different methods of selective pressure according to the arrangement of the community.

The work of ([33]) introduces an alternate method of adaptive assignment strategy (applying a selection operator) called adaptive persecution method (adaptive assignment rules are alternatives to self-adaption). The proposed method is compared to the coincidence probability approach in a non-stationary environment. The experimental results show the superior performance of the proposed method.

Within a Learning Classifier System that uses a diffuse GA ([16]) uses the Intelligent Exploration Method (IEM). This method attempts an adequate exploration according to certain information that is kept of the performance of the environmental agents. The IEM also tries to distinguish diverse phases in the life cycle of the classifier: initial, middle and final.

[32] proposes an interesting operator of adaptive variation which exploits the binary structure of the individuals and presents synergy between breeding and mutation, an approach that applies to the Multi-Objective Optimization (MOO).

[37] offers an algorithm called Bee-GA, applied to the well-known problem of the travelling salesman, which tries to establish the balance between a wide diversity of travels and the exploitation of the best ones.

The work of ([27]) examines an approach where different individuals have different associated mutation rates. Individuals in a present population are ranked according to their aptitude and the mutation rate is increased with an individual ranking; the underlying idea is that the good individuals will produce good and close children while the bad individuals will allow the exploration of other regions in the search space.

[1] uses a very particular approach based on a cell-type GA, similar in approach to ([2]), but with a tridimensional community. Said approach, based on a diversity measurement, attempts to achieve that the algorithm, in a gradual way, refines selective pressure, modifying the parameters, specifically the selection probability.

Due to the proven performance of auto-adaption of the mutation steps in continuous possession and with binary codification, the work of ([31]) examines a method of adapting the election of the mutation operator in execution time, but in permutation type codes. The algorithms are evaluated in the TSP domain.

In ([35]) the selection of adaptive operators in a local search context is addressed. Since diversity is a key concept of the EAs, they are considered to be a related idea: the similarity between a candidate solution and the solutions in a search path. That concept and the quality of the solution are used to evaluate the performance of each operator. That paper introduces a new measurement for usefulness for local search operators, which is compared to a measurement based on the Pareto-Dominancia.

From the above mentioned it can be extracted, obvious from the beginning, that there is a marked tendency in the authors, from the beginning to contemporary ones, to resolve the EED in a direct way by manipulating, whether it be in static or dynamic ways, the breeding and mutation operators to refine them and that the representation of real values (typically used in MOO) have been able to take the possibilities of this resolution approach a little farther. In the same way, the special separation (within the Cellular GA) is a timely subject that has contributed more recently to the EED.

A recent non-standard EA approach is the Learnable Evolution Model (LEM, ([22]). The LEM presents a new idea to employ techniques of Automatic Learning to guide the generation of new generations. The LEM proposal executes an Evolutionary Algorithm in a "Learning Mode". Such a mode considers why certain individuals in a population are superior in relation to others in performing certain tasks. The reasons are expressed as inductive hypothesis and are used, once established, to create new individuals. A very important aspect of LEM is that they can present "evolutionary leaps" in which the aptitude value shows marked improvement, which apparently is due to the discovery of the correct sense of the direction of evolution; due to this behavior the LEM can be considered a type of "Intelligent Evolution". The central aspect of the evolution in the LEM is the Learning Mode, which creates new populations using hypothesis about the best individuals found in past populations. Specifically, the Learning Mode consists of two processes: the generation of hypothesis, in which it is determined, characterizing the differences between individuals of high aptitude and individuals of low aptitude according to past generations, and the process of the hypothesis, which generates new individuals on the basis of the learned hypothesis. In this way new individuals are produced, not just based on the semi-random operations of a classic EA, but using a reasoning process where hypothesis are generated and processed concerning individuals in the population.

The other non-standard approach that is presented is that of the Estimation of Distribution Algorithms (EDA) ([26]). An EDA is a method based principally on substituting breeding and mutation by estimating and later sampling a learned probability distribution from the selected individuals. This group of algorithms has been the object of attention by the scientific community involving evolutionary computer studies and probabilistic graphic models. The population of individuals is reseeded in each generation according to the probability distribution obtained from the best individuals in the previous generation. Precisely due to the fact that the population is not obtained from individuals, but from the obtained probability distributions, an EDA can have the luxury of disregarding the use of breeding and mutation operators. The approximation of an EDA consists of a probabilistic heuristic search based on an EA and on three basic steps, to iterate:

- select some individuals of the population
- estimate the probabilistic model of those selected individuals
- sample the learned probability distribution in order to obtain a new population of individuals

Once the previous approaches are expounded, it can be very clearly seen how these last two are openly different from the general tendency of the EAs, and therefore considered non-standard, and how their way of approaching the EED differs from frequently used approaches. For information purposes, there is a doctoral thesis ([24]), that in a way mixes the two approaches, that is to say, it tries to apply the learning approach (used in LEM) through Bayesian classifiers within the framework of an EDA.

The behavior of the LEM in relation to the EED is rather unusual because a type of progressive partition of the search space is performed, since at the beginning the program that establishes the learning rules generalizes the rules in order to explore, and assures that there are no unchecked areas where an optimum could be found; later this same program is applied to specializing, as much as possible, the rules found.

Without a doubt the way in which the EED can be resolved from an EDA is related to the power of expression of the probabilistic model used, the way this has been learned and how to sample. The same matter, within the LEM approach, is very similar. Because it must be seen that the learning model is so effective that it allows me to construct hypothesis separating the good population from the bad, and therefore they are later used to generate new individuals backing up those hypothesis. We can see therefore the following questions arise:

- How great is the model's power of expression (whether it be probabilistic or learning) used to generate new individuals?
- Is that model being adequately applied (whether it be sampling or processed)? In other words: Am I seeing improvements in each generation, solutions that lead to a true optimum state?

Interestingly, the work that relates to the non-standard approaches with the EED is an attempt to try to involve the known crossover and mutation operators again, since originally these non-standard approaches were sub-estimated or disregarded. The work of ([21]) is one of the first to disseminate that tendency, involving the mutation operator within an EDA called UMDA (univariate marginal distribution algorithm) and at the same time involving the mutation in another EDA called FDA (factorized distribution algorithm), the results showing that while the algorithm takes more time to converge (increased generations), the size of the needed population is less. ([29]), using an adaptive mutation operator in the MT-FDA (Mixture of Trees factorized distribution algorithm) shows that there is improvement in the EED. ([15]) also uses two mutation operators in the framework of an EDA; that is, he uses a typical interchange operator of bit value and also a mutation operator since it is related to the model of probability that is generated by the EDA. The article concludes that the second operator (the one involving the probability model) improves the search process.

# III. Conclusions

It could be concluded that in work related to EAs to palliate the EED, the traditionally focused on the following aspects:

Manipulation, in static or dynamic form, of breeding and mutation operators to refine them to the extent achieved in the search is in a way a very intuitive concept; trying to explore the beginning and exploit the end, the crux of the matter has been related to how to determine when to begin or end an initial or final phase within the search process.

The representation with real values (especially those applied in MOO) have provided a huge amount of related work and in the same way the EAs that imply spatial separation, in more recent times, has provided a new way of approaching the EED.

The non-standard approaches presented, the LEM and the EDAs, if in their conception do not involve breeding and mutation operators, have recently opted to involve mutation operators, especially when they are in harmony with the probabilistic model used, that have shown improvements in the search process.

There is a clear absence of contributions that try to confront the EED from the essence of the non-standard approaches presented, which is to thoroughly review the expressive power of the models used in generating new individuals in each generation and the way in which those individuals are created; here they can resume matters previously presented confronting the EED:

- How big is the power of expression of the model (whether it be probabilistic or learning) that is used to generate new individuals?
- Is that model being adequately applied (whether it be by sampling or requesting) In other words, am I obtaining improvements in each generation, solutions that are leading to a true optimum state?

It can be clearly seen that all future work that tries to respond to the previous matter is going to shed more light on EED in the EAs.

## References

**[1]** Al-Naqi, A.T. Erdogan, T. Arslan, Balancing Exploration and Exploitation in Adaptive Three-Dimensional Cellular Genetic Algorithm via Probabilistic Selection Operator, 2010 NASA/ESA *Conference on Adaptive Hardware and Systems* (AHS-2010), Anaheim, California, USA, June 15-18, 2010.

**[2]** Alba, E., Dorronsoro, B. The Exploration/Exploitation Tradeoff in Dynamic Cellular Evolutionary Algorithms, *IEEE Transactions on Evolutionary Computation*, *9(2),* pp 126-142. Abril de 2005. ISSN: 1089-778X.

[3] Bäck, T., Schütz, M., 1996. Intelligent mutation rate control in canonical genetic algorithms. In: Proceedings of the International Symposium on Methodologies for Intelligent Systems, pp. 158–167.

[4] Bäck, T., Fogel, D.B., and Michalewicz, Z. Eds., *Handbook of Evolutionary Computation*, Oxford University.Press, 1997.

[5] Bäck, T., 1997. Mutation parameters. In: Bäck, T., Fogel, D., Michalewicz, Z. (Eds.), *Handbook of Evolutionary Algorithms.* Institute of Physics Publishing Ltd./ Oxford University Press, Bristol/New York, pp. E1.2.1–E1.2.7.

[6] Ballester, J., P., Carter, N J., 2004. An Effective Real-Parameter Genetic Algorithm forMultimodal Optimization. *ACDM conference (April-04, Bristol, UK*). *Adaptive Computing in Design and Manufacture VI.* I.C. Parmee (Ed.), Springer, pp.359-364.

[7] Berry, D.A., Fristedt, B. (1985), *Bandit problems: Sequential allocation of experiments, Monographs on Statistics and Applied Probability*, London: Chapman & Hall.

[8] Davis, L., 1989. Adapting operator probabilities in genetic algorithms. *In: Proceedings of the Third International Conference on Genetic Algorithms*, pp. 61–69.

[9] Deb K., Agrawal R.B. (1995), *Simulated binary crossover for continuous search space*, *Complex System* 9 115-148.

[10] De Jong, K.A. *Evolutionary Computation*, MIT Press, 2002.

[11] Eiben a, A.E.,and Smith, J.E. *Introduction to Evolutionary Computing*, Springer, 2003.

[12] Eshelman, L. J. (1991). The chc adaptive search algorithm: How to have safe search when engaging in nontraditional genetic recombination. In Rawlins, G. J. E., editor, Foundations of *Genetic Algorithms*-1, pages 265-283. Morgan Kauffman.

[13] Eshelman, L.J.,, Schaffer, J.D. Real-Coded Genetic Algorithms and Interval-Schemata. FOGA 1992: 187-202.

[14] Fogarty, T., 1989. Varying the probability of mutation in the genetic algorithm. In: Proceedings of the Third International Conference on Genetic Algorithms, pp. 104–109.

[15] Handa, H. The effectiveness of mutation operation in the case of Estimation of Distribution Algorithms, Biosystems, Volume 87, Issues 2–3, February 2007, Pages 243-251, ISSN 0303-2647

[16] Hamzeh, A., Rahmani, A., Parsa, N. 2006, Intelligent Exploration Method to Adapt Exploration Rate in XCS, Based on Adaptive Fuzzy Genetic Algorithm, *IEEE Conference onCybernetics and Intelligent Systems*, pp:1 - 6.

[17] Hesser, J., Männer, R., 1991. Toward an optimal mutation probability for genetic algorithms. *In: Proceedings of the First International Conference on Parallel Problem Solving from Nature* (PPSN I), pp. 23–32.

[18] Hinterding, R., Michalewicz, Z., Eiben, A.E., 1997. Adaptation in evolutionary computation: *A survey. In: Proceedings of the Fourth IEEE International Conference on Evolutionary Computation*, pp. 65–69.

[19] Hisashi Handa, The effectiveness of mutation operation in the case of Estimation of Distribution Algorithms, *Biosystems, Volume 87,* Issues 2-3, February 2007, Pages 243-251.

[20] Janikow, C.Z., Michalewicz, Z., 1991. An experimental comparison of binary and floating point representations in genetic algorithms. In: *Proceedings of the 4th International Conference on Genetic Algorithms*, pp. 151–157.

[21] Mahnig, T., Muhlenbein, H. 2001. Optimal Mutation Rate Using Bayesian Priors for Estimation of Distribution Algorithms. In *Proceedings of the International Symposium on Stochastic Algorithms: Foundations and Applications* (SAGA '01), Kathleen Steinhfel (Ed.). Springer-Verlag, London, UK, UK, 33-48.

[22] Michalski. R.S. (2000), Learnable Evolution Model: Evolutionary Processes Guided by Machine Learning, *Machine Learning*, 38, pp. 9-40.

[23] Mitchell, M. *An Introduction to Genetic Algorithms*, MIT Press, Cambridge, MA, 1996.

[24] Miquelez. T.M. Avances en algoritmos de estimación de distribuciones. Alternativas en el aprendizaje y representación de problemas. *Tesis Doctoral*. Universidad del País Vasco. 2010.

[25] Mühlenbein, H., Schlierkamp-Voosen, D. 1993. Predictive models for the breeder genetic algorithm i. continuous parameter optimization. *Evolutionary. Computation*. 1, 1 (March 1993), pp. 25-49.

[26] Mühlenbein, H., Paass, G. From Recombination of Genes to the Estimation of Distributions I. Binary Parameters. *PPSN 1996*: 178-187.

[27] Oliveto, P.S., Lehre, P.K., Neumann, F. Theoretical analysis of rank-based mutation -combining exploration and exploitation. In Proceedings of *IEEE Congress on Evolutionary Computation* '2009. pp.1455-1462

[28] Rechenberg, I., Evolutionsstrategie: Optimierung technischer Systeme und Prinzipien der biologischen Evolution , Frommann-Holzboog, Stuttgart, 1973

[29] Santana, R. An analysis of the performance of the Mixture of Trees Factorized Distribution Algorithm when priors and adaptive learning are used. *Technical Report* ICIMAF 2002-180, Institute of Cybernetics, Mathematics and Physics, Havana, Cuba, March 2002.

[30] Schaffer, J.D., Morishima, A. An adaptive crossover distribution mechanism for genetic algorithms. In: *Proceedings of the Second International Conference on Genetic Algorithms*, pp. 36–40.

[31] Serpell, M., Smith, J.E. Self-Adaptation of Mutation Operator and Probability for Permutation Representations in Genetic Algorithms. *Evolutionary Computation* 18(3): 491-514 (2010).

[32] Tan, K.C. & Chiam, S.C. & Mamun, A.A. & Goh, C.K., 2009. "Balancing exploration and exploitation with adaptive variation for evolutionary multi-objective optimization," European *Journal of Operational Research*, Elsevier, vol. 197(2), pages 701-713, September.

[33] Thierens, D. 2005. An adaptive pursuit strategy for allocating operator probabilities. In *Proceedings of the 2005 conference on Genetic and evolutionary computation* (GECCO '05), Hans-Georg Beyer (Ed.). ACM, New York, NY, USA, 1539-1546.

[34] Vekaria, K., Clack, C., 1998. Selective crossover in genetic algorithms: An empirical study. In: *Proceedings of the Fifth International Conference on Parallel Problem Solving from Nature* (PPSN V), pp. 438–447.

[35] Veerapen, N., Maturana, J., and Saubion, F. 2012. An exploration-exploitation compromise based adaptive operator selection for local search. *In Proceedings of the fourteenth international conference on Genetic and evolutionary computation conference* (GECCO '12), Terence Soule (Ed.). ACM, New York, NY, USA, 1277-1284.

[36] White, T., Oppacher, F., 1994. Adaptive crossover using automata. *In: Proceedings of the Third Conference on Parallel Problem Solving from Nature* (PPSN III), pp. 229–238.

[37] Zhao, G., Luo, W., Nie, H., & Li, C. (2008). A Genetic Algorithm Balancing Exploration and Exploitation for the Travelling Salesman Problem. *2008 Fourth International Conference on Natural Computation*, 505-509.