

## Towards an intelligent genetic search: Defining measures of convergence

Roberto Santana\*

\* Center of Mathematics and Theoretical Physics, ICIMAF,  
Calle 15 e/ C y D, Vedado, Havana, Cuba  
Rsantana@cidet.icmf.inf.cu

The ideal of designing a robust and efficient Genetic Algorithms (GAs), easy to use and applicable to a wide range of problems has deserved a lot of research in Evolutionary Computation. Nevertheless a common criticism made to GA is that in order to perform well for a specific application it needs a lot of tuning and parameter fiddling. Several attempts have been made to investigate the influence of different parameters in the performance of the GA. Recently Harik and Lobo[1] have introduced a Parameter-less Genetic Algorithm which tries 'to make life easier to users' by eliminating the need of parameters in a crossover based GA. Their GA gets rid of the selection rate and crossover probability by establishing fixed setting for these parameters.

We believe that one way leading to the creation of robust and efficient evolutionary algorithms passes by the conception of very flexible methods able to adapt their behavior to the nature of the search space. These algorithms would have a wide arsenal of search operators and a group of general decision rules that would allow them to apply a combination of their operators and parameters. The expertise could not be excluded of the design process, tasks like the coding of solutions would still be his responsibility.

Before to go on with our explanation a generalization is needed. GAs belong to the class of Population Based Search Methods that use Selection (PBSMS). Recently another kind of evolutionary algorithms has been incorporated to this class, they are the Estimation Distribution Algorithms (EDA)[4]. Instead of applying the mutation and crossover operators to the selected population, EDA estimate probability distributions of individuals and use this information to generate new points. By making a parametric or structural learning of the dependency relations between the variables[5][6][7], EDA implicitly adapt their search operators to the characteristics of the problem.

The approach we discuss here is in line with techniques that change or adapt the parameter values as the search progresses, nevertheless it exhibits two main differences with previous work on this topic. First, the analysis has been thought to be applied not only to GA but also to other PBSMS. Second, adaptation is achieved by considering rules able to change the application of different operators along the search, and not only by adapting the parameters. For reasons of space we concentrate here in the question of defining measures that could allow to a PBSMS to receive a feedback about its own behavior, and use this information in the next step of the algorithm to improve the search.

In order to define these measures we will utilize the concepts of Exploration and Exploitation, these concepts have been extensively treated in Evolutionary Computation. A good search method is supposed to be able of combining both tasks. We will also use the mathematical framework applied before to the Breeder Genetic Algorithm (BGA) [3] which was designed according to the theories and methods used in the science of livestock breeding. BGA uses research done in Population Genetics, a field which deals with populations of individuals and describes the properties of traits in terms of their means and their degree of variation in the population. The concept of Response to Selection will be the base for our analysis. Let  $\bar{f}(t)$  be the average fitness of the population at generation t. Then the Response to Selection (R) is defined as:

$$R(t) = \bar{f}(t+1) - \bar{f}(t)$$

The amount of selection is measured by selection differential.

$$S(t) = \bar{f}_s(t) - \bar{f}(t)$$

where  $\bar{f}_s(t)$  is the average fitness of the selected parents. The equation for the Response to Selection relates R and S by :

$$R(t) = b(t) \cdot S(t)$$

b(t) us called the realized heritability. For many fitness functions and selection schemes the selection differential can be expressed as a function of the phenotypic standard deviation  $\sigma$ . For truncation selection (selecting the T·N best individuals from a population of size N) one obtains.

$$S(t)/\sigma(t) = I$$

R is a general measure of the improvements obtained in the average fitness of the population by the application of genetic operators, nevertheless this measure keeps hidden a set of relevant characteristics that could help to understand the PBSMS dynamics. For instance, R could increase without the discovering of new points with a better fitness. On the other hand the discovery of new points may not cause an increment in the R. The R equation relates the change in a population's fitness to the standard deviation of its fitness, as well as to the parameters selection intensity and realized heritability . Nevertheless the standard deviation of fitness measures the fitness diversity of the population but not the quality of solutions. Although the realized heritability can give

an estimation of how similar are the offspring to their parents, this measure is difficult to interpret. We introduce the following approach to cope with these problems.

Let  $[F_{\min}^s(t), F_{\max}^s(t)]$  be the interval for the values of the fitness function in the selection set  $S$ . Using the interval  $[F_{\min}^s(t), F_{\max}^s(t)]$  we will cluster all the vectors of the generation  $t$  in three regions determined by the fitness values of the points they contain. In some way these regions attend to group genotypes by their "quality". The regions are:

Region of promising points ( $R\pi$ ),  $R\pi = x / f(x) > F_{\max}^s$   
 Region of exploitation ( $R\epsilon$ ),  $R\epsilon = x / F_{\min}^s \leq f(x) \leq F_{\max}^s$   
 Region of bad points ( $R\beta$ ),  $R\beta = x / f(x) < F_{\min}^s$ .

Let  $\bar{f}_{R\beta}(t) = \sum_{x/x \in R\beta} f(x) \cdot P(x)$ , where  $P(x)$  is the probability of vector  $x$  in the population at the moment  $t$ .

in a similar way we calculate  $\bar{f}_{R\pi}(t)$  and  $\bar{f}_{R\epsilon}(t)$ .

Then  $\bar{f}(t)$  can be expressed as:  $\bar{f}(t) = \bar{f}_{R\beta}(t) + \bar{f}_{R\pi}(t) + \bar{f}_{R\epsilon}(t)$ .

$\bar{f}_{R\beta}(t)$ ,  $\bar{f}_{R\pi}(t)$  and  $\bar{f}_{R\epsilon}(t)$  are the average fitness contributions of genotypes in each different region to  $\bar{f}$ .

Instead of using  $F_{\min}^s(t)$  and  $F_{\max}^s(t)$  we can use arbitrary values  $f\beta$  and  $f\pi$  ( $f\beta < f\pi$ ) to group the genotypes in regions according to their fitness. We will conveniently cluster genotypes in generation  $t+1$  using  $F_{\min}^s(t)$  and  $F_{\max}^s(t)$  values instead  $F_{\min}^s(t+1)$  and  $F_{\max}^s(t+1)$ .  $R$  can be then decomposed as follows:

$R(t) = \bar{f}(t+1) - \bar{f}(t) = (\bar{f}_{R\beta}(t+1) - \bar{f}_{R\beta}(t)) + (\bar{f}_{R\pi}(t+1) - \bar{f}_{R\pi}(t)) + (\bar{f}_{R\epsilon}(t+1) - \bar{f}_{R\epsilon}(t))$

In this paper we have obtained a decomposition of  $R$  that let us to express the contribution of points in the different regions to the improvements achieved by the population based search algorithm. Figure 1 shows the evolution of  $R$  for the optimization of a univariate function with 100 variables using a Univariate Estimation Distribution Algorithm[4]. Although  $R$  is always positive, the exploration of new points provokes that in some generations  $R\beta$  reaches negative values. Figure 2 shows the corresponding values  $F_{\min}^s(t)$ ,  $F_{\max}^s(t)$  and the average in the interval  $[F_{\min}^s(t), F_{\max}^s(t)]$  for each generation. The use of the decomposition of  $R$  allows to compare different genetic operators taking into account the way they contribute to a positive  $R$ . The difference between the contributions of each regions to  $R$  can be used during the evolution to adapt the operators and parameters looking for a better exploitation or exploration. As we stated before, the question of defining the decision rules that relate the  $R$  values with the change of GA's operators and parameters is beyond the scope of this paper. Diverse theoretical and practical issues have to be examined in order to establish such rules, for instance the impact of different selection operators and the structure of the fitness landscapes[2] in the characteristics of genotypes with values in  $[F_{\min}^s(t), F_{\max}^s(t)]$ .

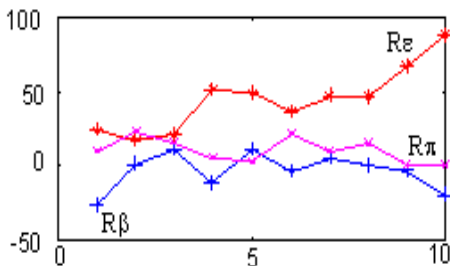


Figure 1

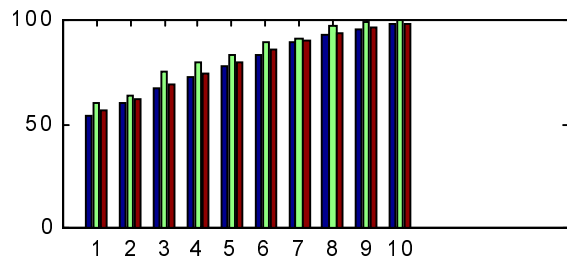


Figure 2

## References

- [1] Harik, G. & Lobo F. (1999). A parameter-less genetic algorithm. Illegal Report 99010. January 1999.
- [2] Manderic k, B., de W egner, M. & Spiessens,P. (1991). The Genetic Algorithm and the Structure of the Fitness Landscape. In Belew,R.K. & Booker, L. editors, Proceedings of the Fourth International Conference on Genetic Algorithms, pp. 45{52, San Mateo:Morgan Kaufman.
- [3] Mühlenbein, H. & Schlierkamp-Voosen D. (1994). The science of breeding and its application to the breeder genetic algorithm. Evolutionary Computation 1, pp. 335--360.
- [4] Mühlenbein, H. (1998). The equation for response to selection and its use for prediction, Evolutionary Computation 5, pp. 303--346.
- [5] Mühlenbein, H.; Mahnig, T., Ochoa A. (1998). Schemata, Distributions and Graphical Models in Evolutionary Optimization, to appear in Journal of Heuristics Vol 5, No. 2. Also at <http://set.gmd.de/AS/ga.publineu.html#index>.
- [6] Ochoa A., Soto M., Santana R., Madera J. C., Jorge N. (1999). The Factorized Distribution Algorithm and The Junction Tree: A Learning Perspective. Proceedings of the II Second Symposium on Artificial Intelligence. ISAS' 99
- [7] Pelikan, M.; Goldberg D. E., Cantú-Paz E.(1998). Linkage problem, Distribution estimation and Bayesian Estimation Networks. Illegal Report 98013. November 1998.