History and Immortality in Evolutionary Computation

Benoit Leblanc[1], Evelyne Lutton[1], Bertrand Braunschweig[2], and Hervé Toulhoat[2]

[1] INRIA, projet FRACTALES 78150 Le Chesnay - France
{Benoit.Leblanc, Evelyne.Lutton}@inria.fr
[2] Institut Français du Pétrole 1 et 4, avenue de Bois-Préau
BP 311 - 92852 Rueil-Malmaison Cedex - France
{Bertrand.Braunschweig, Herve.Toulhoat}@ifp.fr

Abstract. When considering noisy fitness functions for some CPU-time consuming applications, a trade-off problem arise: how to reduce the influence of the noise while not increasing too much computation time. In this paper, we propose and experiment some new strategies based on an exploitation of historical information on the algorithm evolution, and a non-generational evolutionary algorithm.

1 Introduction

Handling noise in Evolution Algorithms has already been studied for the reason that most real-world problems present some noisy behavior, with many possible origins. This difficulty has often been successfully overcome by raising the population size [3] or by making multiple evaluations of the same individual ([5], [9]), using an average as fitness score.

We address here problems where fitness is noisy and in the same time computationally expensive, reducing the applicability of the previous solutions. Considering that each fitness evaluation bears important information that we do not want to lose, an exploitation of the *history* of evaluations is a solution to reduce the misleading noise.

Such a technique has already been experimented by Sano and Kita in [10] for noisy functions, by Corn and al [2] and Zitzler and al [11] for multiobjective optimisation. In Section 2, we propose a similar system of history-based fitness scoring, relying on a *genetic database*. Then in Section 3, it is shown that this genetic database may also be used to produce offspring. A sharing technique complements this scheme, it is described in Section 4. Finally experiments on two multimodal test functions are presented.

2 Historical Information

2.1 Motivation

Inspired by the principles of Darwinian evolution, evolutionary algorithms (EA) are based on the concept of evolving population. The important size of population guarantees the redundancy of information (genes and their expression) and

P. Collet et al. (Eds.): EA 2001, LNCS 2310, pp. 128–139, 2002.

its diversity, so the "death" of old individuals is not a problem, but is rather seen as an important evolution mechanism.

Here we deal with the class of problems where the total number of individuals created during the evolution is limited. This constraint arises for example when the fitness evaluation takes a long time. Moreover if the evaluation is subject to noise, the problem of accuracy of information becomes crucial. As stated before, we cannot afford raising too much the population size or the number of evaluations for the same individual.

To reduce the effect of noise, we therefore propose to use similarities between individuals (many instances of a single individual frequently coexists inside a population). Going further in this direction, we may also consider the whole information produced along the evolution: it often happens that an individual is a copy – or a slightly disturbed copy – of a "dead" ancestor. As we will see below, keeping track of all evaluations performed along the evolution provide another way to reduce the noise of the fitness function.

Moreover, if we can use a metric on the search space that makes sense (i.e. on which we can define a regularity property such as: two individuals that are similar with respect to this metric have similar fitness values), the previous idea may be extended. This implies that we assume some regularity properties of the underlying signal. This is a common hypothesis for many "denoising" techniques in signal analysis [6]. Fitness evaluations may be then averaged for individuals that lie in a given neighbourhood (with appropriate weights, related to the fitness regularity assumption). The resulting computation time overload for the EA remains negligible in the case of time consuming fitness.

2.2 An Implementation for Real-Coded Genomes

Sano and Kita [10] proposed to use the history of search to refine the estimated fitness values of an individual, using the fitness evaluations of individuals similar to it. Their approach is based on a stochastic model of the fitness function that allows to use a maximum likelihood technique for the estimation of the underlying fitness function.

Here we make the assumption that the underlying fitness function is regular with respect to the search space metric. Let us first define:

- The search domain:

$$S = \prod_{i=1}^{m}[a_i, b_i], \text{ with } \forall i \in \{1, ..., m\} \ (a_i, b_i) \in \mathbb{R}^2 \text{ and } \mathrm{a_i} < \mathrm{b_i} \tag{1}$$

- A *max* distance on S:

$$\forall x, y \in S, \quad d_\infty(x, y) = max_{i \in \{1,..,m\}} \left(\frac{|x_i - y_i|}{b_i - a_i} \right) \tag{2}$$

  The divider $(b_i - a_i)$ ensures that each component of a vector has the same weight in the distance regardless of the extent of its domain.