Genetic Algorithms and Evolution Strategies: Similarities and Differences

Frank Hoffmeister frank@unido.uucp

Thomas Bäck baeck@unido.uucp

University of Dortmund, Dept. of Computer Science XI
P.O. Box 500 500, D-4600 Dortmund 50, FRG

Abstract

Evolution Strategies (ESs) and Genetic Algorithms (GAs) are compared in a formal as well as in an experimental way. It is shown, that both are identical with respect to their major working scheme, but nevertheless they exhibit significant differences with respect to the details of the selection scheme, the amount of the genetic representation and, especially, the self-adaptation of strategy parameters.

1 Introduction

Genetic Algorithms (GAs) and *Evolution Strategies* (ESs) are two strata of consciously pursued attempts to imitate principles of organic evolution processes as rules for optimum seeking procedures. Both rely upon the collective learning paradigm gleaned from natural evolution and implement the principles 'population', 'mutation', 'recombination' and 'selection'. In addition, ESs try to use a collective self–learning mechanism to adapt its strategy parameters during the optimum search (*adaptive search*).

The aim of this paper is to show the similarities and differences between both approaches in a formal as well as in an experimental way by viewing at some test functions from Schwefel [Sch75] and Rudolph [Rud90]. In particular, the paper is only an excerpt of a larger investigation on the same subject [HB90].

2 Evolution Strategies

Evolution Strategies emerged from Rechenberg's work in the late sixties, which resulted in a (1+1)–ES with a deterministic step size control [Rec73]. For two model functions Rechenberg was able to show its convergence and the achieved rate of convergence. Schwefel extended the (1+1)–ES towards a $(\mu+\lambda)$–ES and (μ,λ)–ES by applying principles from organic evolution more rigorously. As a result Schwefel proposed an ES capable of *self-adapting* some of its strategy parameters [Sch77, Sch81a]. Born proposed also a population based (μ+1)–ES with the additional concept of a *Genetic Load*, for which he proved the convergence with probability 1 [Bor78]. From the different ES variants Schwefel's ESs will be presented, since it is closest to organic evolution and best suited for the later comparison with GAs.

Schwefel distinguishes between a (μ+λ)–ES and a (μ,λ)–ES. Both ESs fit into the same formal framework with the only difference being the limited life time of individuals in case of (μ,λ)–ES. Thus, only a formal description of the (μ,λ)–ES is presented:

$$(\mu,\lambda)\text{–ES} \quad = \quad (P^0, \mu, \lambda, r, m, s, \Delta\sigma, \Delta\theta, f, g, t) \tag{1}$$

where

$$\begin{array}{lcllr}
P^0 & = & (a_1^0, \ldots, a_\mu^0) \in I^\mu & I = \mathbf{R}^n \times \mathbf{R}^n \times \mathbf{R}^w & \text{population} \\
\mu & \in & \mathbf{N} & & \text{number of parents} \\
\lambda & \in & \mathbf{N} & \lambda \geq \mu & \text{number of offspring} \\
r & : & I^\mu \to I & & \text{recombination operator} \\
m & : & I \to I & & \text{mutation operator} \\
s & : & I^\lambda \to I^\mu & & \text{selection operator} \\
\Delta\sigma & \in & \mathbf{R} & & \text{step-size meta-control} \\
\Delta\theta & \in & \mathbf{R} & & \text{correlation meta-control} \\
f & : & \mathbf{R}^n \to \mathbf{R} & & \text{objective function} \\
g_j & : & \mathbf{R}^n \to \mathbf{R} & j \in \{1, \ldots, q\} & \text{constraint functions} \\
t & : & I^\mu \to \{0,1\} & & \text{termination criterion}
\end{array}$$

Each individual $a_i^t = (x^t, \sigma^t, \theta^t) \in P^t$ within a population consists of three vectors, namely, the set of *object variables* $x^t \in \mathbf{R}^n$, the set of *step sizes* $\sigma^t \in \mathbf{R}^n$, and the set of *inclination angles* $\theta^t \in \mathbf{R}^w$. σ^t and θ^t are internal strategy parameters which control the mutation of x^t.

According to the traditional point of view a population P^t consists of μ parents which produce an intermediate population P'^t made up from λ offspring by means of the recombination operator r and the mutation operator m:

$$\begin{array}{lcll}
P'^t & = & (a_1'^t, \ldots, a_\lambda'^t) & \\
a_i'^t & = & m(r(P^t)) & ; \forall i \in \{1, \ldots, \lambda\}
\end{array} \tag{2}$$

In case of a (μ+λ)–ES the intermediate population P'^t also consists of the μ parents and the selection operator s is modified to $s : I^{\mu+\lambda} \to I^\mu$. By means of selection P'^t is reduced to μ individuals, which become the set of parents P^{t+1} of the next generation.

$$P^{t+1} \;=\; s(P'^t) \quad ; \forall\, a_i^{t+1} = (x, \sigma, \theta) \;\nexists a_j'^t = (x', \sigma', \theta') : f(x') < f(x) \tag{3}$$

According to the rule 'survival of the fittest individuals' the selection operator s removes the least fit individuals from the intermediate population. Figure 3 illustrates