An Approach to the Study of Sensitivity for a Class of Genetic Algorithms

Márton E. Balázs*
Babeş-Bolyai University
Department of Mathematics and Informatics
str. M. Kogălniceanu 1, 3400 Cluj
Romania

Abstract

This paper defines a notion of sensitivity to the convexity of the fitness function used for the class of genetic algorithms using proportional selection and gives some results relative to this notion. The basic idea is that the convexity of the fitness function influences the selective pressure, a property which has been successfully used in genetic algorithms using fitness scaling but without a theoretical foundation. The results in the paper are extensions of the work of Grefenstette and Baker (Grefenstette and Baker, 1989) and of the results of the author (Balázs, 1993b and Balázs, 1994).

1 INTRODUCTION

In (Balázs, 1993b) we presented a preliminary version of our results concerning the study of sensitivity for a class of genetic algorithms to the convexity of the fitness functions used. This research was suggested by the paper of Grefenstette and Baker (Grefenstette and Baker, 1989) where they showed the impact of the fitness functions on the behaviour of genetic algorithms. In the present paper we give further results on this topic.

Section 2 gives a short presentation of the results in (Grefenstette and Baker, 1989) concerning genetic algorithms using a monotonic selection algorithm and a monotonic fitness function. In Section 3 we prove some results concerning genetic algorithms using uniform selection and convex or concave fitness functions. Based on these results in Section 4 definitions of sensitivity for the considered class of genetic algorithms is given, together with

*Presently at Worcester Polytechnic Institute, Dept. of Computer Science, 100 Institute Rd, Worcester, MA 01609

suffucient conditions for this property of a genetic algorithm. Finally Section 5 gives means to compare the sensitivity of genetic algorithms of the same class and shows how experimental observations on the behaviour of some genetic algorithms can be explained in terms of the results obtained. The final section draws some short conclusions on the paper and points out the direction for the future work on the topic.

2 GENETIC ALGORITHMS USING A MONOTONIC SELECTION ALGORITHM AND A MONOTONIC FITNESS FUNCTION

Selection is probably the most important step in a genetic algorithm because it determines which individuals will contribute (and to what amount) to the creation of a new population. As such the nature of the ***selection method*** used greatly influences the behaviour of a genetic algorithm. As in (Grefenstette and Baker, 1989) we shall consider selection to be partitioned into two steps:

1. the ***selection algorithm***, which assigns to each individual x in the population $P(t)$ a real number $tsr(x,t)$ called its ***target sampling rate***;

2. the ***sampling algorithm***, which given the target sampling rate for each individual in $P(t)$ bulids a new population $P(t+1)$ by creating copies of individuals depending on their target sampling rates.

In the present paper we shall concentrate on the selection algorithm, considering (as in (Grefenstette and Baker, 1989)) that an optimal sampling algorithm is given. Thus we shall study methods for computing the target sampling rate for an individual as well as the effect of such methods on the allocation of trials to hyperplanes. This latter will be characterized for a given hyperplane H, at a moment t by the ***target sampling rate of the hyperplane H*** defined by

$$tsr(H,t) = \sum_{x \epsilon H} \frac{tsr(x,t)}{n(H,t)},$$

where $n(H,t)$ denotes the number of representatives of hyperplane H in $P(t)$.

For an optimization problem formulated in terms of an objective function f the target sampling rate for an individual x at moment t is given as a composition of two functions: a ***selection function*** and a ***fitness function***, that is

$$tsr(x,t) = s(u(x,t),t).$$

The best known selection algorithm is ***proportional selection*** defined by

$$tsr(x,t) = \frac{u(x,t)}{\overline{u}(t)},$$

where $\overline{u}(t)$ denotes the average fitness of the individuals in the population at moment t. The first characterization of genetic algorithms given by Holland in (Holland, 1992) is based on proportional selection.

In situations when the objective function is to be minimized and/or it can take on negative values the fitness function is usually obtained by a transformation of the objective function $u(x,t) = k(f(x),t)$ such that

(i) $u(x,t) \geq 0$ for all x and t;

(ii) for any (fixed) t and for all $x_1, x_2 \epsilon P(t)$ such that $f(x_1) \leq f(x_2)$ the following holds:

- $u(x_1,t) \leq u(x_2,t)$ if f is to be maximized;
- $u(x_1,t) \geq u(x_2,t)$ if f is to be minimized.

These would suggest that the only reason for introducing the fitness function is to transform every optimization problem into a maximization problem. However we shall see later in this paper that the choice of the fitness function greatly influences the behaviour of a genetic algorithm.

In the rest of this paper we shall put one more condition on the fitness function, namely that for any (fixed) t and any $x \epsilon P(t)$

$$0 \leq u(x,t) \leq 1,$$

such that if $f(x_m)$ is the worst objective function value for elements in $P(t)$[1], then $u(x_m,t) = 0$, respectively if $f(x_M)$ is the best objective function value in $p(t)$, then $u(x_M,t) = 1$. Note that this condition is added only to keep our discussion simple, and the only effect it can have is on the values produced as entries for the sampling algorithm, but this can be corrected by a linear transformation on u.

In the following we shall omit the parameter t whenever our discussion refers to a certain (fixed) population. We shall use the full notations each time when the value of t is relevant.

Grefenstette and Baker (1989) showed that even for very simple fitness functions (linear ones) the interpretation of Holland's Schema Theorem is not clear. According to them this problem is due to the fact that the theorem refers to the fitness function, which is a design parameter of the genetic algorithm, instead of giving a characterization in terms of the objective function. They suggest that characterization of genetic algorithms should state "**`how the space defined by the objective function is searched by the genetic algorithm`**".

In the rest of this section we give a short presentation of the results in (Grefenstette and Baker, 1989). To make the discussion simple in the following we shall consider (without loss of generality) that f is to be maximized.

Definition 2.1. A fitness function u is *monotonic* if the following condition holds:

$$u(x) \leq u(y) \text{ iff } f(x) \leq f(y).$$

Note 2.1. As shown in (Grefenstette and Baker, 1989) the class of monotonic fitness functions includes many frequently used fitness functions, thus it constitutes a natural object of study.

Definition 2.2. A selection algorithm is *monotonic* if at any moment t it assigns target

[1] i.e. the lowest in the case of a maximization problem, respectively the highest in the case of a minimization problem

sampling rates to individuals such that

$$tsr(x) \leq tsr(y) \text{ iff } u(x) \leq u(y).$$

Note 2.2. Proportional selection, selection by ranking, as well as many other selection algorithms frequently used in practice are monotonic.

The results presented in the following will refer to genetic algorithms using a monotonic selection algorithm and a monotonic fitness function. To give the central characterization of this section one more definition is needed.

Definition 2.3. Given the population $P(t)$ we say that the hyperplane H_1 is dominated by the hyperplane H_2 in $P(t)$ and denote it by $H_1 \leq_{D,t} H_2$ if

$$\max\{f(x) \mid x \epsilon H_1 \cap P(t)\} \leq \min\{f(x) \mid x \epsilon H_2 \cap P(t)\}.$$

Now we can give the following result((Grefenstette and Baker, 1989)):

Theorem 2.1. In any genetic algorithm using a monotonic selection algorithm and a monotonic fitness function, for any hyperplanes H_1, H_2 in $P(t)$

$$H_1 \leq_{D,t} H_2 \Rightarrow tsr(H_1) \leq tsr(H_2).$$

The proof of the theorem is based on the above definitions and is straightforward.

This theorem, although it gives a weaker characterization of the considered class of genetic algorithms then the Schema Theorem does, is more appealing because it is formulated in terms of the objective function.

A more general characterization of the considered class of genetic algorithms can be given based on the following definition:

Definition 2.4. A hyperplane H_1 is *completely dominated* by another hyperplane H_2, and is denoted by $H_1 \leq_D H_2$ if

$$\max\{f(x) \mid x \epsilon H_1\} \leq \min\{f(x) \mid x \epsilon H_2\}.$$

The following corollary holds:

Corollary 2.1. In any genetic algorithm using a monotonic selection algorithm and a monotonic fitness function, for any hyperplabnes H_1 and H_2

$$H_1 \leq_D H_2 \Rightarrow \forall(t) tsr(H_1, t) \leq tsr(H_2, t).$$

This result states that under the given condition H_2 grows at least as fast as H_1 does in any generation, in any genetic algoritm of the considered class.

In the end of their paper Grefenstette and Baker suggest the search for conditions which allow characterizations of the sensitivity of selection algorithms. The next section we lay the basis for the discussion of sensitivity.

3 GENETIC ALGORITHMS USING PROPORTIONAL SELECTION AND THE CONVEXITY OF FITNESS FUNCTIONS

In the previous section a characterization of genetic algorithms using a monotonic selection algorithm and a monotonic fitness function was given. This covers a broad class of genetic algorithms used in practice. However these results can neither give bounds on the growth of the number of representatives of hyperplanes, nor capture sensitivity aspects of genetic algorithms.

We claim that one possible reason for these deficiencies is that in the class considered some selection algorithms may "work against" some qualities of certain fitness functions, a fact which is not captured by the given characterizations. To illustrate this let us consider that our problem is to maximize the function f_0. Let $m_0 = \min\{f_0(x) \mid P(t)\}$ and $M_0 = \max\{f_0(x) \mid P(t)\}$ for a certain t, and suppose that $0 < m_0 < M_0$. Let us further consider the fitness function u_0 given by

$$u_0(x) = \ln(\frac{f(x) - m_0}{M_0 - m_0} \cdot (e - 1) + 1),$$

which is similar to $u(x) = b - log(f(x))$ considered in (Grefenstette and Baker, 1989).

This fitness function has some nice properties which make it appealing, such as

a. it reduces the danger of premature convergence by damping out differences between large values of $f(x)$;

b. it exaggerates differences between small values of $f(x)$ which has the effect of making the algorithm "sensitive" to small differences between low valued individuals.

Since u_0 is monotonic, for a genetic algorithm using a monotonic selection algorithm and u_0 as fitness function we can apply Theorem 2.1. However there are monotonic selection algorithms which void the mentioned qualities of u_0. Let the target sampling rate of an individual x be defined by

$$tsr(x) = \frac{(e^{u(x)} - 1)\frac{M_0 - m_0}{e-1} + m_0}{\mu}$$

where μ is the mean value of the function values for the individuals in $P(t)$. Obviously this selection algorithm is monotonic, and thus it satisfies the condition of Theorem 2.1. Substituting u_0 for u we obtain

$$tsr_o(x) = \frac{f(x)}{\mu(t)}$$

which means a target sampling rate proportional to the function value. Thus the genetic algorithm using tsr_0 as selection algorithm uses a proportional selection which lacks the mentioned properties.

For the reason mentioned and illustrated above we shall discuss properties of the kind listed for u_0 for a narrower class of genetic algorithms, the ones using proportional selection. Obviously this selection algorithm is monotonic and in addition it has the following property:

For any hyperplane H, at any moment t

$$tsr(H) = \sum_{x \epsilon H} \frac{tsr(x)}{n(H)} = \sum_{x \epsilon H} \frac{u(x)}{\overline{u} \cdot n(H)} = \frac{u(H)}{\overline{u}}$$

where $u(H)$ is the average fitness of the representatives of H in $P(t)$. We shall call such a generic algorithm ***Proportional Selection Genetic Algorithm*** (PSGA).

Note 3.1. Obviously all the results in the previous section are valid for PSGA's using a monotonic fitness function.

Let us now try to give a characterization of PSGA's using a monotonic fitness function with properties similar to those of u_0 above. In order to do this we need some more definitions.

First recall that given a real-valued function $g : D(\subseteq \Re) \rightarrow \Re$ the divided difference of g with respect to the points $x, y \epsilon D$ is defined by

$$[x, y; g] = \frac{g(x) - g(y)}{x - y}.$$

Divided differences of higher order are defined recursively, e.g. the divided difference of g with respect to the points $x, y, z \epsilon D$ is defined as

$$[x, y, z; g] = \frac{[x, y; g] - [y, z; g]}{x - z}.$$

Definition 3.1. We say that u is a *convex fitness function* if for any individuals x_1, x_2 and x_3 in the domain of f

$$[f(x_1), f(x_2), f(x_3); u] \geq 0$$

where [x,y,z;h] denotes the second order divided difference of the function h in the points x, y and z.

The fitness function u is said to be *concave* if for any individuals x_1, x_2 and x_3 in the domain of f

$$[f(x_1), f(x_2), f(x_3); u] \leq 0.$$

The following two theorems give a property of PSGA's using a monotonic and convex, respectively concave fitness functions, which makes them interesting from the point of view of study of sensitivity.

Theorem 3.1. For a PSGA using a monotonic and convex fitness function u, for any x_1, x_2 and x_3 in $P(t)$ such that $f(x_1) < f(x_2) < f(x_3)$

$$f(x_3) - f(x_2) \geq f(x_2) - f(x_1) \Rightarrow tsr(x_3) - tsr(x_2) \geq tsr(x_2) - tsr(x_1).$$

The corresponding result for concave fitness functions is given by

Theorem 3.2. For a PSGA using a monotonic and concave fitness function u, for any x_1, x_2 and x_3 in $P(t)$ such that $f(x_1) < f(x_2) < f(x_3)$

$$f(x_3) - f(x_2) \leq f(x_2) - f(x_1) \Rightarrow tsr(x_3, t) - tsr(x_2, t) \leq tsr(x_2, t) - tsr(x_1, t).$$

Since the proofs of the two theorems are absolutely analogous we shall only give the proof for the second one.

Proof. (Theorem 3.2.) Since u is concave, we have

$$[f(x_1), f(x_2), f(x_3); u] \leq 0$$

which by the definition of the divided difference is

$$\frac{\frac{u(x_3)-u(x_2)}{f(x_3)-f(x_2)} - \frac{u(x_2)-u(x_1)}{f(x_2)-f(x_1)}}{f(x_3) - f(x_1)} \leq 0.$$

Using the hypothesis that $f(x_1) < f(x_2) < f(x_3)$ we obtain

$$\frac{u(x_3) - u(x_2)}{f(x_3) - f(x_2)} \leq \frac{u(x_2) - u(x_1)}{f(x_2) - f(x_1)},$$

which is equivalent to

$$u(x_3) - u(x_2) \leq (u(x_2) - u(x_1)) \frac{f(x_3) - f(x_2)}{f(x_2) - f(x_1)}.$$

The condition $f(x_3) - f(x_2) \leq f(x_2) - f(x_1)$ means that

$$(0 <) \frac{f(x_3) - f(x_2)}{f(x_2) - f(x_1)} \leq 1,$$

by which the previous inequality becomes

$$u(x_3) - u(x_2) \leq u(x_2) - u(x_1).$$

Dividing this latter inequality by $\overline{u}$ we obtain

$$tsr(x_3) - tsr(x_2) \leq tsr(x_2) - tsr(x_1),$$

which completes the proof. □

Although these results are more general the following two corollaries explain better the importance of the above theorems.

Corollary 3.1. For a PSGA using a monotonic and convex fitness function u, for any x_1, x_2 and x_3 in $P(t)$ such that $f(x_1) < f(x_2) < f(x_3)$ and $f(x_3) - f(x_2) = f(x_2) - f(x_1)$

$$tsr(x_3) - tsr(x_2) \geq tsr(x_2) - tsr(x_1).$$

Corollary 3.2. For a PSGA using a monotonic and concave fitness function u, for any x_1, x_2 and x_3 in $P(t)$ such that $f(x_1) < f(x_2) < f(x_3)$ and $f(x_3) - f(x_2) = f(x_2) - f(x_1)$

$$tsr(x_3) - tsr(x_2) \leq tsr(x_2) - tsr(x_1).$$

The proofs of these two corollaries are identical to those of the corresponding two theorems by noting that $f(x_3) - f(x_2) = f(x_2) - f(x_1)$ implies

$$\frac{f(x_3) - f(x_2)}{f(x_2) - f(x_1)} = 1$$

The first corollary states that in a PSGA using a monotonic and convex fitness function the ***distinction between two individuals***, made by their target sampling rates is the greater the higher their objective values are. Stated otherwise, such an algorithm strongly favors the individuals with higher objective values.

On the other hand the second corollary states that a PSGA using a monotonic and concave fitness function damps out the differences between large objective values (exactly as u_0 did).

Although it is not relevant for the further discussions on sensitivity, let us give a characterization of how the convexity (concavity) of the fitness function used influences the allocation of trials to hyperplanes by a PSGA.

We shall only give the result and make the corresponding proof for concave fitness functions, the ones for the convex case being absolutely analogous.

We have the following theorem:

Theorem 3.3. In a PSGA using a monotonic and concave fitness function for any hyperplanes H_1, H_2 and H_3 represented in $P(t)$ such that $H_1 <_{D,t}< H_2 <_{D,t} H_3$ and $f(x_3) - f(x_2) \leq f(x_2) - f(x_1)$ for all $x_1 \epsilon H_1$, $x_2 \epsilon H_2$ and $x_3 \epsilon H_3$

$$tsr(H_3) - tsr(H_2) \leq tsr(H_2) - tsr(H_1).$$

Proof. Let x_i^{min} denote an individual such that $f(x_i^{min}) = \min\{f(x) \mid x \epsilon H_i \cap P(t)\}$ for $i = 1, 2, 3$ and x_i^{max} an individual such that $f(x_i^{max}) = \max\{f(x) \mid x \epsilon H_i \cap P(t)\}$ for $i = 1, 2, 3$.

Using these notations the inequality $f(x_3) - f(x_2) \leq f(x_2) - f(x_1)$ in the hypothesis of the theorem implies

$$f(x_3^{max}) - f(x_2^{min}) \leq f(x_2^{min}) - f(x_1^{max}),$$

which by Theorem 3.2. gives

$$u(x_3^{max}) - u(x_2^{min}) \leq u(x_2^{min}) - u(x_1^{max}).$$

This is equivalent to

$$\sum_{i=1}^{n(H_3)} \frac{u(x_3^{max})}{n(H_3)} - \sum_{i=1}^{n(H_2)} \frac{u(x_2^{min})}{n(H_2)} \leq \sum_{i=1}^{n(H_2)} \frac{u(x_2^{min})}{n(H_2)} - \sum_{i=1}^{n(H_1)} \frac{u(x_1^{max})}{n(H_1)}.$$

Since u is monotonic we have

$$\sum_{i=1}^{n(H_3)} \frac{u(x)}{n(H_3)} - \sum_{i=1}^{n(H_2)} \frac{u(x)}{n(H_2)} \leq \sum_{i=1}^{n(H_2)} \frac{u(x)}{n(H_2)} - \sum_{i=1}^{n(H_1)} \frac{u(x)}{n(H_1)},$$

that is

$$u(H_3) - u(H_2) \leq u(H_2) - u(H_1),$$

which divided by $\overline{u}(> 0)$ gives

$$tsr(H_3) - tsr(H_2) \leq tsr(H_2) - tsr(H_1).$$

Thus our proof is complete. □

Note 3.2. We can make similar notes concerning the behaviour of PSGA's using monotonic and convex (concave) fitness functions relative to the allocation of hyperplanes as we did after Corollary 3.1. and Corollary 3.2.

In the following section we shall use these results to study some sensitivity properties of PSGA's.

4 SENSITIVITY OF A PSGA TO THE CONVEXITY OF THE USED FITNESS FUNCTION

In this section we shall introduce a concept of sensitivity which will make it possible to describe the behaviour of certain PSGA's. Let us first give an intuitive description of what we shall mean by "sensitivity" of a PSGA.

Selection in genetic algorithms is done based on the comparison of the objective function values of individuals in a population. However the distinction between individuals is not directly based on their objective function values but is made via their target sampling rate. As we have seen in Section 2 this is a composition of two functions. In the case of PSGA's one of these function (the selection algorithm) is fixed. It should be intuitively clear that by changing the fitness function one can influence the way a PSGA distinguishes between individuals (this fact has actually been recognized and used in various applications by the means of parametrized fitness functions). In the following we shall try to give a means to measure the way a PSGA distinguishes between individuals (sensitivity) and prove some results on how some characteristics of fitness functions influence this property of the algorithm.

Definition 4.1. In a genetic algorithm using a selection algorithm defined by tsr, by the ***absolute distance*** of two individuals x_1 and x_2 in $P(t)$, denotes the distance between their target sampling rates, i.e.

$$\delta_{tsr}(x_1, x_2) = \mid tsr(x_1) - tsr(x_2) \mid .$$

We define the ***relative distance*** of the two individuals by

$$\rho_{tsr}(x_1, x_2) = \frac{\delta_{tsr}(x_1, x_2)}{\mid f(x_1) - f(x_2) \mid}$$

In the following while speaking about a single algorithm we shall omit the subscript tsr.

With this notion we can give a first definition concerning sensitivity.

Definition 4.2. Let us consider a genetic algorithm and let x_1, x_2, x_3 and x_4 be individuals in some population $P(t)$. We say that the genetic algorithm makes a ***stronger (weaker) distinction*** between x_3 and x_4 than between x_1 and x_2 if

$$\rho(x_3, x_4) \geq \rho(x_1, x_2).$$

Similarly the genetic algorithm makes a ***weaker distinction*** between x_3 and x_4 than between x_1 and x_2 if

$$\rho(x_3, x_4) \leq \rho(x_1, x_2).$$

This definition captures how the selection in a genetic algorithms is influenced by the differences between objective function values.

Let us now see how the convexity (concavity) of a fitness function influences the distinction made by a PSGA between individuals in a population. To simplify the discussion we shall prove the following lemma:

Lema 4.1. If a fitness function is convex (concave) then it is monotonic.

Proof. We first consider the proof for the convex case. We have to prove that if a fitness function u is convex then for any x and y in $P(t)$

$$u(x) \leq u(y) \Leftrightarrow f(x) \leq f(y).$$

Let x_m denote an element in $P(t)$ such that $f(x_m) = \min\{f(x) \mid x \epsilon P(t)\}$ and let us assume that x and y are two individuals in $P(t)$ such that $f(x_m) < f(x) < f(y)$. By the definition of convexity we have

$$[f(x_m), f(x), f(y); u] \geq 0,$$

that is

$$\frac{\frac{u(x_m)-u(x)}{f(x_m)-f(x)} - \frac{u(x)-u(y)}{f(x)-f(y)}}{f(x_m)-f(y)} \geq 0.$$

By the condition imposed on fitness functions in Section 2[2] $u(x_m) = 0$, which together with our hypothesis $f(x_m) < f(x) < f(y)$ makes the previous inequality equivalent to

$$u(x) - u(y) \leq (f(x) - f(y)) \cdot \frac{-u(x)}{f(x_m) - f(x)}.$$

But the right side of this inequality represents a negative number, which means

$$u(x) \leq u(y).$$

The inverse implication is absolute identical except for the rearrangement of the inequation. Thus our proof for the convex case is complete.

For the concave case the proof is similar, except only for that the third individual considered for building the divided difference will be an x_M such that $f(x_M) = \max\{f(x) \mid x \epsilon P(t)\}$.Now the proof of the lemma is complete. □

Note 4.1. The above lemma gives us in the following the possibility to use "convex (concave) fitness function" instead of "monotonic and convex (concave) fitness function".

Let us now give definitions of two types of sensitivity for genetic algorithms.

Definition 4.3. We say that a genetic algorithm is ***high-value sensitive*** if for any $x_1, x_2, x_3 \epsilon P(t)$ such that $f(x_1) < f(x_2) < f(x_3)$ it makes a stronger distiction between x_2 and x_3 than between x_1 and x_2, i.e.

$$\rho(x_2, x_3) \geq \rho(x_1, x_2).$$

Similarly we say that a genetic algorithm is ***low-value sensitive*** if for any $x_1, x_2, x_3 \epsilon P(t)$ such that $f(x_1) < f(x_2) < f(x_3)$ it makes a weaker distiction between x_2 and x_3 than between x_1 and x_2, i.e.

$$\rho(x_2, x_3) \leq \rho(x_1, x_2).$$

By this definition we call high-value sensitive those genetic algorithms which make a stronger distinction in favor of individuals with higher objective function values, while low-value sensitive will be those genetic algorithms which distingush more strongly between individuals having low objective function values.

Theorem 4.1. A PSGA using a convex fitness function is high-value sensitive.

Theorem 4.2. A PSGA using a concave fitness function is high-value sensitive.

Proof. We only give the proof for Theorem 4.1 (i.e. the convex case), the one for Theorem 4.2 being similar.

[2] i.e. that the value of the fitness function corresponding to the "worst" objective function value is 0.

Let x_1, x_2 and x_3 be three individuals in $P(t)$ such that $f(x_1) < f(x_2) < f(x_3)$. The convexity of the fitness function (let it be u) means

$$\frac{\frac{u(x_3)-u(x_2)}{f(x_3)-f(x_2)} - \frac{u(x_2)-u(x_2)}{f(x_2)-f(x_1)}}{f(x_3) - f(x_1)} \geq 0,$$

which by the hypothesis is equivalent to

$$\frac{u(x_3) - u(x_2)}{f(x_3) - f(x_2)} - \frac{u(x_2) - u(x_2)}{f(x_2) - f(x_1)} \geq 0.$$

Dividing this inequality by $\overline{u}$ ve have

$$\frac{tsr(x_3) - tsr(x_2)}{f(x_3) - f(x_2)} - \frac{tsr(x_2) - tsr(x_2)}{f(x_2) - f(x_1)} \geq 0,$$

and this is exactly what we wanted to prove, i.e.

$$\rho(x_2, x_3) \geq (\leq)\rho(x_1, x_2).$$

□

This theorem explains us the behaviour in selection of a certain class of genetic algorithms (PSGA's). It makes clear for instance what makes a PSGA using fitness function u_0 defined in Section 3 have the listed properties. More precisely, the function $u : [m_0, M_0] \to [0, 1]$ defined by

$$u(t) = \ln(\frac{t - m_0}{M_0 - m_0} \cdot (e - 1) + 1)$$

is twice (in fact infinite times) derivable on (m_0, M_0) and has its second order derivative negative for all $t \epsilon (m_0, M_0)$, i.e. is concave (in analytical sense). But this means that its second order divided difference is negative for any three points in its domain and so for any three distinct values of the objective function. According to the above theorem this means that a PSGA using u_0 as fitness is low-value sensitive.

Let us consider another example. Filipič (1992) described a GA application where the fitness function used was similar to the following (except that he defined it for a minimization problem):

$$u_k(x) = \left(\frac{f(x) - f_{min}}{f_{max} - f_{min}} \right)^{1/k},$$

where f_{min} and f_{max} are (estimated) lower, respectively upper bounds of the objective function f and $k > 1$ is a parameter of the algorithm. He (correctly) noted that by using this fitness functions "...only the best solutions ... receive distinguished high fitness values." Following the same reasoning as above we can easily show that the considered fitness function is convex (for any $k > 1$) which, by Theorem 4.1 explains the behaviour noted by Filipič (1992).

However the above conclusion is not the only reason for which we chose the latter example. The note made in (Filipic, 1992) concerning the effect of the fitness function u_k on the behaviour of the genetic algorithm also has a first part, namely "Increasing k, while $k > 1$, causes a wider and wider range of high cost solutions[3] to be assigned low fitness...". What is interesting about this note is that it gives (even if just qualitatively) a comparison between the effect of several fitness functions ("...more and more..."). In the next section we shall try to formally capture this phenomenon by defining a way to compare the effects of two different fitness functions of a certain class on the behaviour of a PSGA.

5 COMPARING THE SENSITIVITIES OF PSGA -S

As we did throughout this paper in this section too we shall only concentrate on a certain class of fitness functions. Let us start by defining the class to be considered.

Definition 5.1. We say that a fitness function is ***uniform*** if it is either convex or concave.

In the following we shall only consider PSGA's using a uniform fitness function. We call such a genetic algorithm ***Uniform*** PSGA (UPSGA).

We shall now define an ordering relation over the class of UPSGA's which allows comparing two such algorithms with respect to their sensitivity.

Definition 5.2. Let A_1 and A_2 two UPSGA's using the fitness functions u_1 and u_2 respectively. We say that A_1 is ***less sensitive*** than A_2 and denote this relation by $A_1 \leq_S A_2$ if for any population $P(t)$ and any individuals x_1, x_2 and x_3 in $P(t)$ such that $f(x_1) < f(x_2) < f(x_3)$

$$\rho_{tsr_1}(x_2, x_3) - \rho_{tsr_1}(x_1, x_2) \leq \rho_{tsr_2}(x_2, x_3) - \rho_{tsr_2}(x_1, x_2).$$

We also say that A_2 is ***more sensitive*** than A_1.

Put in words this definition says that a UPSGA is more sensitive than another one if it makes a stronger distinction between higher objective function values than the second one.

The central theorem of this section gives a sufficient condition for an UPSGA to be more sensitive than another.

Theorem 5.1. Let A_1 and A_2 be two UPSGA's using the fitness functions u_1 and u_2 respectively. If for any population $P(t)$ and any individuals x_1, x_2 and x_3 in $P(t)$ such that $f(x_1) < f(x_2) < f(x_3)$ the inequality

$$[f(x_1), f(x_2), f(x_3); u_1] \leq [f(x_1), f(x_2), f(x_3); u_2]$$

is true, then $A_1 \leq_S A_2$.

To prove this theorem we shall first prove the following lemma:

Lemma 5.1. Let A_1 and A_2 two UPSGA's using the fitness functions u_1 and u_2 respectively. If for any population $P(t)$ and any individuals x_1, x_2 and x_3 in $P(t)$ such that $f(x_1) < f(x_2) < f(x_3)$ the inequality

[3] Remember that the problem there was to minimize a function.

$$[f(x_1), f(x_2), f(x_3); u_1] \leq [f(x_1), f(x_2), f(x_3); u_2]$$

is true, then $u_1(x) \geq u_2(x)$ for all $x \epsilon P(t)$.

Proof. Let (as in the case of the Lemma 4.1.) x_m denote a member of $P(t)$ such that $f(x_m) = \min\{f(x) \mid x \epsilon P(t)\}$ and x_M a member of $P(t)$ such that $f(x_M) = \max\{f(x) \mid x \epsilon P(t)\}$. By our conditions on fitness functions $u_1(x_m) = u_2(x_m) = 0$ and $u_1(x_M) = u_2(x_M) = 1$. Let further x be a members of $P(t)$ such that $f(x_m) < f(x) < f(x_M)$. By the hypothesis we have

$$[f(x_m), f(x), f(x_M); u_1] \leq [f(x_m), f(x), f(x_M); u_2],$$

which, by the above assumptions implies

$$\frac{1 - u_1(x)}{f(x_M) - f(x)} - \frac{u_1(x)}{f(x) - f(x_m)} \leq \frac{1 - u_2(x)}{f(x_M) - f(x)} - \frac{u_2(x)}{f(x) - f(x_m)}.$$

By rearranging this inequality we obtain

$$\frac{u_2(x) - u_1(x)}{f(x) - f(x_m)} + \frac{u_2(x) - u_1(x)}{f(x_M) - f(x)} \leq 0,$$

that is

$$(u_2(x) - u_1(x))[\frac{1}{f(x) - f(x_m)} + \frac{1}{f(x_M) - f(x)}] \leq 0.$$

But this inequality can only be true if

$$u_2(x) \leq u_1(x).$$

Thus our proof is complete. $\Box$

Let us now return to Theorem 5.1.

Proof. (Theorem 5.1.) Let x_1, x_2 and x_3 be three individuals in $P(t)$ such that $f(x_1) < f(x_2) < f(x_3)$. By the hypothesis of out theorem the inequality

$$[f(x_1), f(x_2), f(x_3); u_1] \leq [f(x_1), f(x_2), f(x_3); u_2]$$

is true. By the definition of divided differences this means

$$\frac{u_1(x_3) - u_1(x_2)}{f(x_3) - f(x_2)} - \frac{u_1(x_2) - u_1(x_1)}{f(x_2) - f(x_1)} \leq \frac{u_2(x_3) - u_2(x_2)}{f(x_3) - f(x_2)} - \frac{u_2(x_2) - u_2(x_1)}{f(x_2) - f(x_1)}.$$

This is equivalent to

$$\overline{u}_1 \cdot [\frac{tsr_1(x_3) - tsr_1(x_2)}{f(x_3) - f(x_2)} - \frac{tsr_1(x_2) - tsr_1(x_1)}{f(x_2) - f(x_1)}] \leq$$

$$\leq \overline{u}_2 \cdot [\frac{tsr_2(x_3) - tsr_2(x_2)}{f(x_3) - f(x_2)} - \frac{tsr_2(x_2) - tsr_2(x_1)}{f(x_2) - f(x_1)}],$$

or, equivalently

$$\frac{tsr_1(x_3) - tsr_1(x_2)}{f(x_3) - f(x_2)} - \frac{tsr_1(x_2) - tsr_1(x_1)}{f(x_2) - f(x_1)} \leq$$
$$\leq [\frac{tsr_2(x_3) - tsr_2(x_2)}{f(x_3) - f(x_2)} - \frac{tsr_2(x_2) - tsr_2(x_1)}{f(x_2) - f(x_1)}] \cdot \frac{\overline{u}_2}{\overline{u}_1}.$$

But by Lemma 5.1 $u_1(x) \geq u_2(x)$ for all $x \epsilon P(t)$, which implies $\overline{u}_1 \geq \overline{u}_2$, or equivalently $\overline{u}_2/\overline{u}_1 \leq 1$. By this the previuos inequality implies

$$A_1 \leq_S A_2,$$

and our proof is complete. □

Put into words, Theorem 5.1 means that for genetic algorithms with proportional selection using uniform fitness functions, the "more convex" the fitness function, the more sensitive the algorithms are to high objective function values.

Let us now return to the fitness function $u_k(k > 1)$ defined in the previous section. Actually this is a family of fitness functions defined by the parameter k. Let k_1 and k_2 be two real numbers such that $1 < k_1 < k_2$. It can be easily shown that $u''_{k_1}(t) < u''_{k_2}(t)$ for all $t \epsilon (f_{min}, f_{max})$ which implies that the same relation holds between the second order divided differences for any three elements of the same segment[4]. This latter means that the condition of Theorem 5.1 is fulfilled and thus the UPSGA using u_{k_1} as fitness function is more sensitive then the one using u_{k_2}. This conclusion explains the second note on the fitness functions u_k quoted from (Filipič, 1992).

6 CONCLUSIONS

In this paper we introduced ways to study the sensitivity of a class of genetic algorithms. The results presented explain the behaviour in selection of these algorithms and gives theoretical justification for some qualities of certain fitness functions recognized by several authors as well as for the success of using scaling methods for fitness functions.

In (Balázs, 1993b) we also showed that for UPSGA's a lower bound on the growth of the number of representatives of hyperplanes can be given.

Our further work in this direction will concern the study of scaling of fitness functions for a UPSGA. Some preliminary results on this are given in (Balázs, 1994).

[4] u"(t) denotes the second order derivative of the function u in t.

References

Balázs, M.E. (1993a) On an Experiment Using Genetic Algorithms for Solving Equations, ***submitted to Studia Universitatis Babeş-Bolyai.***

Balázs, M.E. (1993b) On a Class of Fitness Functions for Genetic Algorithms Using Proportional Selection, ***submitted to Revue d'Analyse Numerique et de Théorie de l'Approximation.***

Balázs, M.E. (1994) Genetic Algorithms and Applications, ***PhD Thesis***, Babeş-Bolyai University, Cluj.

Filipič, B. (1992) Enhancing Genetic Search to Schedule a Production Unit, Proc. ECAI'92 Neumann, B. Ed., John Wiley & So.

Goldberg, D. (1989) Genetic Algorithms in Search, Optimization and Machine Learning, ***Addison-Wesley.***

Grefenstette, J.J. and Baker,J.E. (1989) How Genetic Algorithms Work: A Critical Look at Implicit Parallelism, ***Proceedings of ICGA'89***, Ed. Schaffer.

Holland, J.H., (1975) Adaptation in Natural and Artificial Systems, ***Ann Arbor: The University of Michigan Press.***