

ON A CLASS OF FITNESS FUNCTIONS FOR GENETIC ALGORITHMS USING PROPORTIONAL SELECTION

MÁRTON-ERNÖ BALÁZS

(Cluj-Napoca)

1. INTRODUCTION

In [2] Grefenstette and Baker discussed the impact of the fitness functions on the behaviour of genetic algorithms. They showed situations where Holland's Schema Theorem (*see* [3]) does not have a clear interpretation and suggested, for a rather large class of genetic algorithms a simple but useful characterization of the implicit parallelism.

In the present paper we study a class of fitness functions for genetic algorithms using proportional selection. The first section contains the results in [2] concerning genetic algorithms using a monotonic fitness function and a monotonic selection algorithm. In the second section we study a class of fitness functions which render certain sensitivity properties to genetic algorithms using proportional selection. The last section presents some conclusions as well as our ideas about possible further work.

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

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 in [2] we shall consider selection to be partitioned into two steps :

- a. the *selection algorithm*, which assigns at each moment t , to each individual x a real number $tsr(x, t)$, called its *target sampling rate*;
- b. the *sampling algorithm*, which builds a new generation by creating copies of individuals depending on the target sampling rate.

In the present paper we shall concentrate on the selection algorithm, considering that an optimal sampling algorithm is given. Thus we shall study rules for computing tsr for an individual as well as the effect of such rules 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

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

where $n(H, t)$ denotes the number of representatives of hyperplane H in the population at moment t (denoted by $P(t)$).

For a problem formulated in terms of an objective function f , the target sampling rate is given by composing two functions: a *fitness function* u and a *selection algorithm* s , that is

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

For the rest of this paper we shall consider that the fitness does not depend on t , i.e. $tsr(x, t) = s(u(x), t)$.

In situations when the objective function is to be minimized or when it can take on negative values the fitness function is obtained by a transformation of the objective function such that $u(x) = k(f(x))$. However these are not the only reasons for using fitness functions. As we shall see, they greatly influence the behaviour of genetic algorithms.

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

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

where $\bar{u}(t)$ denotes the average fitness of the individuals in the population at moment t . The first characterization of genetic algorithms, given by Holland [3] is based on proportional selection.

In [2] Grefenstette and Baker 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 characterizations 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 [2]. To make discussion simple in the followings we shall consider (without loss of generality) that f is to be maximized.

DEFINITION 1.1. A fitness function u is monotonic if the following condition holds:

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

Note 1.1. As shown in [2] this class of monotonic fitness functions include many frequently used fitness functions, thus it constitutes a natural subject of study.

DEFINITION 1.2. A selection algorithm is monotonic if it assigns a target sampling rate to each individual, at any moment such that

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

Note 1.2. Proportional selection, selection by ranking, as well as many other known selection algorithms are monotonic.

The results presented in the followings will concern 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 1.3. Given the population $P(t)$ we say that the hyperplane H_1 is dominated by the hyperplane H_2 , denoting it by $H_1 \leq_{D,H_2} H_2$ if

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

Now we can give the following result ([2]):

THEOREM 1.1. In any genetic algorithm using a monotonic selection algorithm and a monotonic fitness function, for any hyperplanes, H_1, H_2 in $P(t)$

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

The proof of Theorem 1.1 is based on the definitions and is straightforward.

This theorem, although it gives a weaker characterization of the considered class of genetic algorithms than 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 1.4. A hyperplane H_1 is completely dominated by another hyperplane H_2 , denoted by $H_1 \leq_{D,H_2} H_2$ if

$$\max \{f(x) | x \in H_1\} \leq \min \{f(x) | x \in H_2\}.$$

The following corollary holds:

COROLLARY 1.1. In any genetic algorithm using a monotonic selection algorithm and a monotonic fitness function, for any hyperplanes H_1 and H_2

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

This result states that under the given conditions H_2 grows at least as fast as H_1 does in any generation, in any genetic algorithm 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 presents some of our results in this direction.

3. GENETIC ALGORITHMS USING PROPORTIONAL SELECTION

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 prac-

tice. However these results can neither give bounds on the growth of the 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. To illustrate this let us consider the fitness function

$$u_0 = \log(f(x)) - 1, \quad e < f(x) < m \text{ for all possible } x$$

(which is similar to $u = b - \log(f(x))$ considered in [2]) where f is to be maximized.

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. makes a great difference between large and small values of $f(x)$.

For a genetic algorithm using a monotonic selection algorithm and the fitness function u_0 we can apply Theorem 1.1. However there are monotonic selection algorithms which annul the mentioned qualities of u_0 . Let for instance the target sampling rate be defined by

$$tsr_0(x, t) = \frac{e^{u+1}}{M_f(t)},$$

where $M_f(t) = \max\{f(x) | x \in P(t)\}$. Substituting u_0 for u we obtain

$$tsr_0(x, t) = \frac{e^{u_0+1}}{M_f(t)} = \frac{f(x)}{M_f(t)}.$$

The genetic algorithm using tsr_0 and u_0 uses a linear selection which lacks the above-mentioned properties.

For the reason mentioned and illustrated above we shall discuss genetic algorithms which uses proportional selection. Obviously this selection algorithm is monotonic and in addition it has the following property:

For any hyperplane H , at any moment t

$$\begin{aligned} tsr(H, t) &= \sum_{x \in H} \frac{tsr(x, t)}{n(H, t)} \\ &= \sum_{x \in H} \frac{u(x)}{\bar{u}(t)n(H, t)} \\ &= \frac{u(H, t)}{\bar{u}(t)} \end{aligned}$$

where $u(H, t)$ is the average fitness of the representatives of H in $P(t)$. We shall call such genetic algorithms Proportional Selection Genetic Algorithms (PSGA).

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

As shown in [2], results similar to those presented in the previous section can be obtained for strictly monotonic fitness functions and strictly monotonic selection algorithms. In the followings by monotonic fitness function we shall mean *strictly monotonic* ones. Obviously proportional selection is strictly monotonic.

Let us now try to give a characterization for 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.

DEFINITION 3.1. We say that u is a convex fitness function if for any x_1, x_2 and x_3

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

where $[x, y, z; h]$ denotes the second order divided difference of the function h in x_1, x_2, x_3 .

The fitness function u is said to be concave if for any x_1, x_2 and x_3

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

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

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 $f(x_1) < f(x_2) < f(x_3)$

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

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 latter 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, since u is monotonic, 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

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

by which the previous equality becomes

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

Dividing this inequality by $\bar{u}(t)$ we obtain

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

which completes the proof.

Note 3.2. Theorem 3.1. states that a PSGA using a monotonic and convex fitness function stresses the difference between high values of $f(x)$. As such they are prone to premature convergence but seem to speed up growth for the best fit individuals. This property makes this class of genetic algorithms to be appealing in advanced stages of search. On the other hand, according to Theorem 3.2., a PSGA using a monotonic and concave fitness function damps out differences between high values of $f(x)$, reducing by this the rise of premature convergence. Due to this property genetic algorithms of this class are advantageous in early stages of search.

Note 3.3. The previous note gives an intuitive explanation why scaling methods are suggested and used in selection practice.

Note 3.4. We could have obtained similar results by defining convex and concave selection algorithms. The reason why we did not do so is that in our future work we intend to give some means to compare sensitivity of members of a class of genetic algorithms and in our view the first step in this direction would be the study of the sensitivity induced by the fitness function for the same selection algorithm.

In the following we shall study how convexity (concavity) influences the allocation of trials to hyperplanes. In order to do this we shall discuss in details PSGA-s using a monotonic and concave fitness function, the results and all the steps made for the convex case being similar.

We have the following result:

THEOREM 3.3. For a PSGA using a monotonic and concave fitness function and for any hyperplanes H_1, H_2 and H_3 in $P(t)$ such that $H_1 <_{D,t} H_2 <_{D,t} H_3$

$$\begin{aligned} \max\{f(x) | x \in H_3 \cap P(t)\} + \max\{f(x) | x \in H_1 \cap P(t)\} &\leq \\ &\leq 2 \min\{f(x) | x \in H_2 \cap P(t)\} \Rightarrow \\ \Rightarrow tsr(H_3, t) - tsr(H_2, t) &\leq tsr(H_2, t) - tsr(H_1, t) \end{aligned}$$

Proof. We shall denote by x_i^{\min} a member of $P(t)$ such that $f(x_i^{\min}) = \min\{f(x) | x \in H_i \cap P(t)\}$ for $i = 1, 2, 3$, and by x_i^{\max} a member of $P(t)$ such that $f(x_i^{\max}) = \max\{f(x) | x \in H_i \cap P(t)\}$ for $i = 1, 2, 3$.

With these notation the left side of the conclusion of the theorem becomes

$$f(x_3^{\max}) + f(x_1^{\max}) \leq 2f(x_2^{\min}),$$

which is the same as

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

By the concavity of the fitness function u this implies

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

which is equivalent to

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

Since u is monotonic we have

$$\sum_{x \in H_3} \frac{u(x)}{n(H_3, y)} - \sum_{x \in H_2} \frac{u(x)}{n(H_2, y)} \leq \sum_{x \in H_2} \frac{u(x)}{n(H_2, y)} - \sum_{x \in H_1} \frac{u(x)}{n(H_1, y)},$$

that is

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

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

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

Thus our proof is complete.

Note 3.5. Theorem 3.3. and the corresponding result for convex fitness functions give some explanation on the behaviour of a class of genetic algorithms, based on the nature of the underlying fitness function. They establish conditions under which a genetic algorithm makes strong distinction between either very fit individuals, or individuals with low fitness.

Note 3.6. Since, as stated at the beginning of this paper (as well as in [2]), the fitness function is a design parameter of a genetic algorithm these theorems can give some support in choosing this parameter.

However there is one more interesting thing about PSGA-s using a monotonic and either a convex or a concave fitness function, which concerns the relative growth of representatives of two hyperplanes. We have the following theorems.

THEOREM 3.4. For a PSGA using a monotonic and convex fitness function u and for hyperplanes H_1 and H_2 in $P(t)$ such that $H_1 <_{D,t} H_2$ the following inequality is true:

$$tsr(H_2, t) \geq \left[1 + \frac{m_u^{[x_1^{\min}, x_1^{\max}]}}{u(H_1, t)} (f(x_2^{\min}) - f(x_1^{\max})) \right] tsr(H_1, t),$$

where $m_u^{[x_1^{\min}, x_1^{\max}]}$ = $\min \left\{ \frac{u(x) - u(x_1^{\min})}{f(x) - f(x_1^{\min})} \mid x \in [x_1^{\min}, x_1^{\max}] \right\}$.

THEOREM 3.5. For a PSGA using a monotonic and concave fitness function u and for hyperplanes H_1 and H_2 in $P(t)$ such that $H_1 \leq_{n,t} H_2$ the following inequality is true:

$$tsr(H_2, t) \geq \left[1 + \frac{m_u^{(x_2^{\min}, x_2^{\max})}}{u(H_1, t)} (f(x_2^{\min}) - f(x_1^{\max})) \right] tsr(H_1, t),$$

$$\text{where } m_u^{(x_2^{\min}, x_2^{\max})} = \min \left\{ \frac{u(x_2^{\min}) - u(x)}{f(x_1^{\min}) - f(x)} \mid x \in (x_2^{\min}, x_2^{\max}) \right\}.$$

Note 3.7. Since the values of f are considered to be bounded and u is monotonic both of the considered minimums exist.

We give the proof for Theorem 3.5, for the corresponding one with convex fitness function the proof is analogous.

Proof. (Theorem 3.5.) Since u is a concave fitness function, using the same notation as above we have

$[f(x_1^{\max}), f(x_2^{\min}), f(x); u] \leq 0$ for any $x: f(x) \in (f(x_2^{\min}), f(x_2^{\max}))$, that is

$$\frac{u(x) - u(x_2^{\min})}{f(x) - f(x_2^{\min})} - \frac{u(x_2^{\min}) - u(x_1^{\max})}{f(x_2^{\min}) - f(x_1^{\max})} \leq 0 \text{ for any } x: f(x) \in (f(x_2^{\min}), f(x_2^{\max})).$$

This is the same with

$$u(x_2^{\min}) \geq u(x_1^{\max}) + \frac{u(x) - u(x_2^{\min})}{f(x) - f(x_2^{\min})} (f(x_2^{\min}) - f(x_1^{\max})),$$

which implies

$$u(x_2^{\min}) \geq u(x_1^{\max}) + m_u^{(x_2^{\min}, x_2^{\max})} (f(x_2^{\min}) - f(x_1^{\max})).$$

By the relation $H_1 \leq_{n,t} H_2$ this implies

$$tsr(H_2, t) \geq tsr(H_1, t) + \frac{m_u^{(x_2^{\min}, x_2^{\max})} (f(x_2^{\min}) - f(x_1^{\max}))}{\bar{u}(t)}$$

which is the same with

$$\frac{tsr(H_2, t)}{tsr(H_1, t)} \geq 1 + \frac{m_u^{(x_2^{\min}, x_2^{\max})}}{u(H_1, t)} (f(x_2^{\min}) - f(x_1^{\max})).$$

This latter inequality is equivalent to the one we had to prove, i.e.

$$tsr(H_2, t) \geq \left[1 + \frac{m_u^{(x_2^{\min}, x_2^{\max})}}{u(H_1, t)} (f(x_2^{\min}) - f(x_1^{\max})) \right] tsr(H_1, t),$$

Thus for PSGA-s which uses a monotonic and either a convex or a concave fitness function we could give a bound of the relative growth of hyperplanes in $P(t)$.

4. CONCLUSIONS AND FUTURE WORK

The main point of this paper is that there is a rather large class of genetic algorithms for which some sensitivity properties can be established depending on the nature of the underlying fitness function. As in our main reference [2] the results are formulated in terms of the objective function.

The interpretation of the mentioned results show that for PSGA-s using monotonic fitness functions the latter should be chosen to be concave in early stages of operating and convex in the advanced ones. This also suggests that it is useful to scale the fitness function depending on a parameter which somehow reflects the progress of search. Concerning this problem there are at least two questions which may arise:

a. How to design a (parametrized) fitness function to exploit the presented results?

b. What would the parameter which directs scaling be and how should it change during the operation of the genetic algorithm?

In a further paper we shall try to give some possible answers to these two questions.

REFERENCES

1. Balázs, M. E., *On an Experiment for Using Genetic Algorithms for Solving Equations*, submitted to Studia Univ. Babeş-Bolyai (1993).
2. Grefenstette, J. J. and Baker, J. E., *How Genetic Algorithms Work: A Critical Look at Implicit Parallelism*, Proceedings of ICGA'89, Ed. Schaffer (1989).
3. Holland, J. H., *Adaptation in natural and artificial systems*, Ann Arbor: University Michigan Press (1975).

Received 27 VIII 1993

Babeş-Bolyai University
Department of Mathematics and
Informatics
Str. M. Kogălniceanu 1,
3400 Cluj-Napoca,
Romania