On functions with a given fitness–distance relation

To appear in 1999 Congress on Evolutionary Computation. July 13-17, Orlando, Florida

Leila Kallel
CMAP – UMR CNRS 7641,
Ecole Polytechnique,
Palaiseau 91128, France
kallel@cmapx.polytechnique.fr

Bart Naudts
Dept. of Maths & Comp. Sci.
University of Antwerp, RUCA
Groenenborgerlaan 171
B–2020 Belgium
bnaudts@ruca.ua.ac.be

Marc Schoenauer
CMAP – UMR CNRS 7641,
Ecole Polytechnique,
Palaiseau 91128, France
marc@cmapx.polytechnique.fr

Abstract- Recent work stresses the limitations of fitness distance correlation (FDC) as an indicator of landscape difficulty for genetic algorithms (GAs). Realizing that the fitness distance correlation (FDC) value cannot be reliably related to landscape difficulty, we investigate whether an interpretation of the whole correlation plot can yield reliable information about the behavior of the GA. Our approach is as follows. We present a generic method for constructing fitness functions which share the same fitness versus distance-to-optimum relation (FD relation). Special attention is given to FD relations which show no local optimum in the correlation plot, as is the case for the relation induced by Horn's longpath. We give an inventory of different types of GA behavior found within a class of fitness functions with a common correlation plot. We finally show that GA behavior can be very sensitive to small modifications of the fitness–distance relation.

1 Introduction

A number of summary statistics have been proposed to characterize classes of fitness functions with similar GA behavior. We mention fitness distance correlation [JF95], correlation length and operator correlation [MdWS91], epistasis [Dav91], schema variance [RS95] and hyperplane ranking [DWP95]. These summary statistics originated from empirical observations and their efficiency is widely discussed (e.g. [RVSK97, Alt97, QVSG98]).

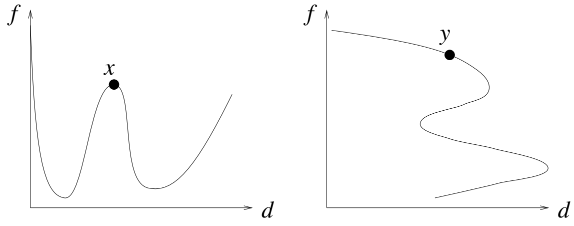
Studies of the above properties and statistics have mostly been limited to the construction of examples and counterexamples. An interesting direction, scarcely taken, is to construct and study classes of fitness functions which share the same property or statistic value, and to understand the differences in GA behavior on these functions.

Examples of for the latter approach are the NK-landscapes [Kau89] which provide a simple way of constructing classes of functions with a given amount of epistasis (order of interaction between genes). The study of NK-landscapes gives useful information about the relation between the structure of the landscape and the order of epistasis. Note that the notion of epistasis is not to be confused with Davidor's epistasis measure, which cannot distinguish between different orders of epistasis (contrary to the Anova tables of [RW95].)

Fitness distance correlation measures the degree to which the fitness values increase with the individuals approaching the optimum in Hamming distance.

Recent work [NL98] shows that the values produced by this statistic are often unrelated to the behavior of the GA. To summarize: unless one restricts its scope to some narrow function class, the only reliable information fitness distance correlation gives about the difficulty or structure of the fitness landscape is the sign of the correlation coefficient value.

Of course, the whole fitness distance plot gives more information than the summary statistic (correlation value). It is shown in [KS97] how the sample lay out in the fitness–distance space and diversity comparisons can be useful in choosing problem-dependent initialization procedures. But further interpretation of this plot can be dangerous, since only partial information about the structure of the fitness landscape is displayed. A local optimum in the fitness landscape is not necessarily visible in the fitness distance plot. For example, the point x in figure (a) below is both a local optimum in the fitness landscape and in the correlation plot. The point y in figure (b) may or may not be a local optimum in the fitness landscape, depending on the distribution of the neighbors of y.

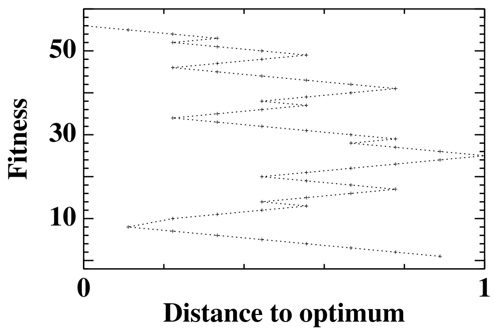


(a) peak (b) no peak

Two types of correlations plots (a) x is both a local optimum in the fitness landscape and in the correlation plot. (b) y could be a local optimum in the fitness landscape.

Hence "continuous" lines in the FDC correlation plot do not necessarily define a feasible path to the optimum of f, although in an intuitive way this may seem to be the case. They may even hide intractable local optima.

In order to study the effects of invisible information in the plot (e.g. the presence of local optima, the number of genes represented by one point, etc.), we will focus on fitness–distance relations which show no local optimum in the correlation plot, as is the case in figure (b). What difference in



(a) Correlation plot

(b) Relation H

Figure 1: (a) Correlation plot of Horn's long path problem. (b) FD Relation H for Horn's long path problem. The points marked on the fitness axis are the extrema $f_1, \ldots, f_K$ of the function H.

GA behavior can we expect for a fixed FD relation?

This paper is organized as follows: Section 2 introduces some conventions and the main results of this section. Section 3 presents a generic method to construct fitness functions which follow a given fitness–distance relation, with Horn's longpath [HG95] as the main example. The method is then used in Section 4 to construct a class of functions whose members share the same FD relation. We make an inventory of the different types of GA behavior observed in this class, and study the effects of small modifications in the FD relation.

2 Main Results

2.1 Experimental conditions

Unless explicitly stated, we use a generational GA with population size of 50, linear ranking selection, uniform crossover at rate 0.5, and $1/l$ mutation with a rate of 0.5 per individual. The stopping criterion is set at 100,000 fitness evaluations without improvement of the best value, with a maximum of 1,000,000 fitness evaluations. All results are averaged over 50 independent trials.

2.2 Notations

we propose the following hierarchy of *convergence quality* of a given GA and fitness function (first introduced in [NL98]):

1. The GA detects the optimum of the fitness function in each trial (with a small[1] variance on convergence time).
2. Each GA trial still detects the optimum of the fitness function, but now the variance on the number of iterations needed for detection is comparatively large.
3. Only some GA trials detect the optimum of the fitness function.
4. The GA does not perform better than random search. It cannot detect the optimum of the fitness function.

2.3 Summary of the results

Section 4 studies a class of fitness functions f_k sharing the same fitness-distance relation of Eq. 8, sketched in Fig. 3. It shows that according to the strings distribution in the plot (choice of k value in Eq. 9), GA convergence can be either stable and straightforward (convergence quality 1), or unstable (convergence quality 2). Moreover, GA trajectories in the correlation plot varies according to the choice of the rule. However, changing the rule yields a consequent change of strings density in the correlation plot, which partially explains the observed differences.

A mostly interesting result is obtained with the function f_1', which follows the same FD relation than f_1, and is equal to f_1 for points with below average distances (to optimum) only. Yet, most GA trials fail to find the optimum of f_1' (convergence quality 3, vs. 2 for f_1). This is due to the presence of an intractable local optimum with respect to *hamming* distance, in the middle of the *upper path* (the dot line with the highest fitness values in the correlation plot) of f_1'.

On the other hand, a small perturbation of the fitness values allows to to get rid of local constantness in the landscape. Applied to f_1 GA convergence quality (initially 2) is 1. Whereas GA behavior is still the same (quality 3) when the perturbation is applied to f_1'.

[1]We consider the variance as small if it has a lower order of magnitude than convergence time.

(a) $k = 1$ (b) $k = 5$ (c) $k = 9$

Figure 2: Number of strings in logarithmic scale versus normalized k-th zero position in $[k, \ell + k - \mathrm{d}(s, s^*)]$, with $\ell = 50$.

3 General Construction

In this section, we propose a way of constructing fitness functions which share the same fitness–distance relation. The shape of this relation is similar to that induced by Horn's long path [HG95] (Fig. 1). Only successive points on Horn's long path are within Hamming distance 1 from each other; they define a unique feasible path for a hill climber.

Let us drop the requirement of proximity of neighboring points on the path, and constrain the relation between the fitness of a string and its distance to the optimum to functions of the form

$$H(\textit{fitness}) = \textit{distance-to-optimum}, \tag{1}$$

with

$$H : [0, f\mathrm{max}] \to \{0, \ldots, \ell\}, \tag{2}$$
$$H^{-1}(0) = \{f\mathrm{max}\} \tag{3}$$

In this case, no local optimum is visible in the correlation plot, and we obtain a long-path shaped FD relation (each fitness value is associated with a unique distance). Note that the choice of the optimum is arbitrary, and unless otherwise mentioned, it is set at $s^* = 11\ldots11$.

All fitness functions satisfying FD relation H can now be described by assigning values to a *rule* function, which maps the strings to natural numbers. *rule*(s) indicates which branch of the plot (at a given distance from optimum) will be chosen as the fitness of string s. It follows

$$\forall s \in E,\ \textit{rule}(s) \in \{1, \ldots, \#H^{-1}(\mathrm{d}(s, s^*))\} \tag{4}$$

A fitness value can be assigned to each string by following the convention that lower *rule* values correspond to lower fitness values, i.e., if $rule(s) = 1$ and $d = \mathrm{d}(s, s^*)$, then the lowest fitness value of the set $H^{-1}(d)$ is assigned to s. If $rule(s) = k$, with $k \leq \#H^{-1}(d)$, then the k-th greatest fitness value of the set $H^{-1}(d)$ is assigned to s.

3.1 Example of Horn's long path rule

As an example, we give the *rule* function for Horn's long path, assuming that we have already chosen an adequate function H, such that the obtained correlation plot is that of the long path (see Fig. 1). We denote the extrema of H by $f_1, f_2, \ldots, f_K$. The optimum of the long path is reached in $110\ldots0$.

Horn, Goldberg and Deb [HG95] give a recursive algorithm which either returns the position of a string in the long path or indicates that this string is off the path. Let *pos*(s) be the position of s in the path. It ranges from 1 to $3.2^{(\ell-1)/2} - 1$, which is the length of the path when ℓ is odd. The special value *pos*$(s) = 0$ indicates that s is off the path.

Armed with this *pos* function, we define the *rule* function by

$$\textit{rule}(s) = \sum_{i=1}^{K} i\,[p(f_i) < \textit{pos}(s) < p(f_{i+1})], \tag{5}$$

where $p(f_i)$ corresponds to the position in the path of the (unique) string at fitness f_i. The brackets are used for the indicator function. The $p(f_i)$ are determined recursively by

$$p(f_1) = 1, \tag{6}$$
$$p(f_{i+1}) = H(f_i) - H(f_{i+1}) + p(f_i) \tag{7}$$

4 Case Study Of A Fitness–Distance Relation

The first part of this section uses the method presented in the previous section to construct a class of fitness functions whose members share the same fitness–distance relation. The properties of some of the fitness landscapes in this class are investigated.

(a) 85% of the GA-trials (b) 15% of the GA-trials (c) possible HC behavior

Figure 3: The GA and hill-climber trajectories (best individual so far for each generation) plotted on the correlation plot for the function f_1 associated with rule r_1. In (a) the GA trajectories go straight up to s^0, then follow the difficult upper path to converge to the optimum. In (b) the GA does not go through the difficult upper path, and converges rather quickly. Three sorts of one-bit-flip hill-climber behavior are shown in (c). Of all the trials, about 50% take the path along the left side to s^*, 25% take that along the right side to s_0, and 25% get rapidly stuck in a local optima inside the cone of H.

$$H(x) = \frac{\ell/2}{x_{\max} - \pi} \left(x \sin(x) + x_{\max} \right), \qquad (8)$$

for $x \in [0, x_{\max}]$, with $x_{\max} = 2\pi\kappa - \pi/2$. After a suitable restriction of its domain, H satisfies the conditions of Eqs. 2 and 3. For the sake of symmetry, and to avoid a bias towards above average distances, the function H is slightly translated to the left, which centers the plot around $\ell/2$ (see Fig. 3). The fittest string $s^* = 11\ldots11$ is at leftmost position. The string represented by the rightmost position of the plot, which is in Hamming distance closest to $00\ldots00$, and on the sin curve, is noted s^0.

Having defined the relation H between fitness values and the distance of the strings to the optimum, we proceed by defining a parametrized class of *rule* functions, denoted $\{r_k\}$, hereby obtaining a class of fitness functions $\{f_k\}$ satisfying the FD relation H.

NOTATIONS: Given a string s, we denote by $Z_k(s)$ the position of the k-th zero of s. As the range of $Z_k(s)$ is $\{k, \ldots, k+\ell-\mathrm{d}(s, s^*)\}$, we call $(Z_k(s)-k)/(\ell-\mathrm{d}(s, s^*)+1)$ the *normalized k-th zero position.*

The rule r_k is defined in such a way that the further the k-th zero is in a string, the higher the fitness of this string is:

$$r_k(s) = \left\lfloor \frac{Z_k(s) - k}{\ell - \mathrm{d}(s, s^*) + 1} \# H^{-1}(\mathrm{d}(s, s^*)) \right\rfloor + 1, \qquad (9)$$

for all strings s and positions $k \in \{1, \ldots, \ell\}$. Since this rule takes values between 1 and $\# H^{-1}(\mathrm{d}(s, s^*))$ inclusive, it follows the conditions of Eq. 4, and therefore, defines a fitness function f_k following the fitness–distance relation H of Eq. 8.

Figure 2 shows the distribution of strings at distance d from the optimum in function of the normalized k-th zero position.

Note that the rule is defined in such a way that the further the k-th zero is in a string, the higher the fitness of this string is. Hence, the plot also characterizes the distribution of strings in the fitness–distance plot. We observe that

- there are more low-fit strings at distance $\ell/2 + a$ than at distance $\ell/2 - a$. Similarly, there are less high-fit strings at distance $\ell/2 + a$ than at distance $\ell/2 - a$;
- if $k = 1$ then, at a given distance from optimum, the number of strings decreases when fitness increases (Fig. 2(a)).
- as k increases, the number of strings with high fitness values increases (Fig. 2(b,c)).
- when $k > (\mathrm{d}(s, s^*)+1)/2$, there are more strings at top than at bottom fitness values (Fig. 2(b) with $\mathrm{d}(s, s^*) = 5$ or Fig. 2(c) with $\mathrm{d}(s, s^*) = 15$).

Note that strings at distances beyond $\ell/2$, with minimal k-th zero position ($Z_k(s) = k$), define an easy path up to s^0. In fact, within these points, the fitness function is determined by the Hamming distance towards a point of s^0. The same thing happens symmetrically: strings at distances shorter than $\ell/2$ define an easy path (determined by the Hamming distance) to s^*. These two paths will be called the *paths along the left-hand and right-hand side* in the next sections. We also define the *upper path* as the points of the correlation plot that link s^0 to s^* via highest fitness values.

4.2 Hill-climber behavior on the function class $\{f_k\}$

The f_k-landscape of the one-bit-mutation hill-climber contains many local optima, while experiments show almost none for the landscape of the GA defined in Sect. 2. Figure 3(c) presents the different kinds of hill-climber trajectories:

1. For $k = 1$, the one-bit-mutation hill-climber and the $1/l$ mutation (1+1)-, (1+50)- and (50+50)-evolution strategies show the same behavior: either they take the easy left-most path to the optimum, or they get stuck in some local optimum (which is often one of the strings in s^0).
2. As k increases (not shown), most HC trials (e.g. 70% when $k = 5$) get stuck in a local optimum within 100 fitness evaluations. On the other hand, the GA and the evolution strategies perform increasingly well.

4.3 GA-behavior on the function class $\{f_k\}$

Let us now concentrate on the behavior of the GA on the functions $\{f_k\}$. Two types of GA-trajectories can be observed in the plots of Fig. 3(a,b):

• Some trials go inside the fitness–distance plot, as in Fig. 3(b). They always find the optimum within 100 generations.

• The other trials follow the right side and then take the upper path from right (s^0) to left (s^*), as in Fig. 3(a). We note that:

1. When $k = 1$, this path is followed by the majority of trials. It takes the GA a very long time (1000 up to 5000 generations) to move from the right side of the upper path to the left side. Convergence is quicker on the left half of the upper path.
2. With k increasing, the convergence time on the upper path drops below 400 generations, but the number of trials which follow this path decreases.

The distribution of strings in the fitness–distance plot partially explains the observed GA-behavior, since the GA is mainly attracted to dense and high-fit regions. We once again distinguish between $k = 1$ and k increasing:

1. When $k = 1$, the dense areas define an easy along the right-hand side to s^0 because of the conic shape of H. Therefore, most GA-trials take this path to s^0. Then, due to elitism, the GA necessarily follows the upper path from s^0 to s^*.
2. With k increasing, the number of high-fit strings increases, and the search on the upper path becomes easier.

The difficulties on the upper path are due to *isolation*. For all points s on the upper path, at a distance $\mathrm{d}(s, s^*) > \ell/2$, the next fitter strings s' are still on the upper path, but at a shorter distance from the optimum. The range of possible values of $Z_k(s')$ is therefore larger than that of $Z_k(s)$. Moreover, given the shape of H, there are probably more fitness values to allocate for s' than for s. A direct consequence of these facts, is that $Z_k(s')$ is likely to be greater than $Z_k(s)$. Hence s and s' are likely to be at a large Hamming distance from each other. In between the strings are low-fit regions.

This picture is reversed for below-average distances from te optimum, i.e., for string s with $\mathrm{d}(s, s^*) < \ell/2$. The number of different fitness values either decreases or remains constant. Each time it decreases, search becomes easier.

We have now shown that different types of GA behavior arise, on the same fitness–distance relation H, according to the distribution of strings over the possible function values (the parameter k tunes this distribution).

In the next sections 4.4 and 4.5, we stick to the case of $k = 1$, and consider two modifications of the rule of Eq. 9. The first modification removes the plateaus in the fitness landscape which are responsible for the problem of isolation described above. The second modification symmetrizes the distribution of strings with respect to the distance $\ell/2$.

4.4 Removing the plateaus with the function f_1^+

The minor modification made to f_1 to obtain f_1^+ is aimed at showing the obvious limitations of any statistical measure of the fitness landscape which is based on fitness values. It also shows the sensitivity of GA's trajectory to local features of the landscape.

We construct f_1^+ by modifying f_1 (which corresponds to rule r_1 defined in Eq. 9):

$$f_1^+(s) = f_1(s) + \sum_{j=2}^{\ell/5} \frac{Z_j(s)}{\ell}\, 10^{-j+1}, \qquad (10)$$

for all strings $s \in \{0,1\}^\ell$. The effect of the additional terms is a small cloud around each point in the fitness–distance plot: the flat areas have been removed by the small differences (less than 10^{-1}) in fitness values. Yet, the macroscopic overall fitness–distance relation remains the same: the density of strings at each fitness level ($\pm 10^{-1}$) are the same. The fitness differences between strings in different fitness–distance positions remain the same, at precision 10^{-1}.

The Hamming landscape of f_1^+ keeps its many local optima. The one-bit-flip hill-climber shows the same three kinds of trajectories as plotted in Fig. 3(c) for f_1.

But evolution strategies behave differently. While the (1+1)-ES finds the upper path of f_1 intractable, it swiftly optimizes the upper path of f_1^+ (convergence within 6000 fitness evaluations). Moreover, only 30% of the trials (vs. 50% for f_1) take the path along the right-hand side, as shown in Figure 4.

(a) H (b) H^c (c) H_0

Figure 5: Typical GA behavior according to the reference function choice for the fitness-distance relation: (a) Function $\dot{H}$(b) The mirror image of H over the vertical $\ell/2$ axis (c) The original sine-based reference function, asymmetric with respect to the vertical $\ell/2$ axis.

The GA can also escape easily the from the absorbing right-hand side of f_1, and goes inside the cone of H.

It is important to note that any statistical *a priori* measure based on fitness values would give roughly the same results for f_1^+ and f_1. Yet, the GA's speed of convergence is completely different for both functions. The upper path, initially intractable for the (1+1)-ES, becomes trivial to optimize for this algorithm. The observed trajectories of GA and evolution strategies are different too: in the case of f_1^+ they go often inside the cone, rather than taking the absorbing path along the right-hand side when confronted with f_1.

4.5 Adding symmetry to the distribution with f_1'

Again we make a small modification to the fitness function f_1, this time by changing the rule r_1. The result is a fitness function which satisfies the fitness–distance relation H, but is intractable for the GA due a too large basin of attraction of a local optimum.

Recall that the rule r_k of Eq. 9 distributes more high-fit strings at distances beyond $\ell/2$ than before $\ell/2$, as can be seen in Fig. 2. To symmetrize the distribution of individuals in the plot with respect to $\ell/2$, we propose an alternative rule for above average distances.

Let us use $O_k(s)$ to denote the position of the k-th one in s. Then we define:

$$r'_k(s) = \begin{cases} \left\lfloor \dfrac{Z_k(s) - k}{\ell - \mathrm{d}(s, s^*) + 1} \# H^{-1}(\mathrm{d}(s, s^*)) \right\rfloor + 1, \\ \qquad\qquad \text{when } \mathrm{d}(s, s^*) \le \ell/2 \\ \\ \left\lfloor \dfrac{O_k(s) - k}{\ell - \mathrm{d}(s, s^*) + 1} \# H^{-1}(\mathrm{d}(s, s^*)) \right\rfloor + 1, \\ \qquad\qquad \text{when } \mathrm{d}(s, s^*) > \ell/2 \end{cases} \tag{11}$$

Associated with the rules $\{r'_k\}$ is a class of fitness functions $\{f'_k\}$, whose members follow the same fitness–distance relation H as the $\{f_k\}$. The functions f'_k and f_k coincide for strings at distance $d < \ell/2$ from optimum.

Due to the conic shape of H, f'_1 still contains an easy path to s^0, but only via strings with minimal k-th one position. Experiments show that 70% of all GA-trials take this easy path to s^0, similarly to the case of f_1. From s^0 they start on the upper path towards the optimum. In the case of f'_1, though, the search stops in the middle of the plot, around strings with $\mathrm{d}(s, s^*) = \ell/2$.

In fact, strings on the upper path with below average distance have their k-th zero at the highest position, and are very close to the string $s_{10} = 11\ldots1100\ldots00$. Strings with an above average distance have their k-th one at the highest position too, and hence are close to the bitwise complement of s_{10}, which is $s_{01} = 00\ldots0011\ldots11$. It is thus not surprising that the trials which start the upper path in s^0 get stuck in some intractable local optimum around $\ell/2$.

It is interesting to note that, while involving a strong local optimum, the upper path exhibits a locally ideal fitness–distance correlation (value -1). On the other hand, the local optimum in the GA landscape does not exist in the landscape of a GA with complementary crossover.

Removing the plateaus helps to transform the upper path of f_1 into a (1+1)-ES-easy function. This is not the case for f'_k: there is no feasible path linking s_{01} to s_{10} via strings with either a high k-th zero position or a high k-th one position. Thus, removing flat areas does not necessarily make fitness landscapes easier.

We also tested a minor changes in the FD relation H of Eq. 8:

1- We replaced the sine function by a cosine to get the mirror image of H over the vertical $\ell/2$ axis as shown in Fig. 5-b. Where for FD relation $H$and function f'_k, most (64%) GA-trials take the path along the right, and get stuck in s_{01}, experiments show that for the mirror of H, denoted H^c, opposite results hold: few (40%) GA-trials take the path along the right side. On the other hand, since the distribution of strings in f_k is asymmetric with respect to $d\ell/2$ (there are more low-fit strings at above average distances), it is not surprising that moving to H^c scarcely changes the GA's behavior on f_k (with $H 70\%$ trials take the right path, versus 50% with H^{sym}).

2- We also tested the original sine function (not centred over the vertical $\ell/2$ axis) as in Fig. 5-c, then only two trials out of fifty follow the right+upper path with function f_1.

Note that the conic shape of H results in two attractors, s^0 and s^*, for the GA, which generally starts at average distances from s^*. One can expect a that a different shape of the FD relation H will completely change the GA's behavior.

A case study of the GA-behavior on functions with a FD relation which gives a better coverage of the fitness–distance plot can be found in [Kal98].

5 Discussion

We presented a case study of GA behavior on a number of functions, which correlation plot shows no local optimum in the fitness–distance space. The overall conclusion is that convergence quality varies from 1 to 4 according to strings distributions (invisible) in the plot.

The fact is that two neighboring points in the correlation plot are not necessarily close in Hamming distance to each other: The extreme case of function f_1' shows how a locally linear relation between fitness and distance to optimum (the upper path) can hide an intractable local optimum in the Hamming space. On the other hand, the same correlation plot may contain easy paths to the optimum eventually resulting in a OneMax like behavior.

However, the path that is chosen by the GA is heavily dependent on the position of the initial population in the correlation plot, as a small translation of the fitness–distance relation along the distance axis completely changes the dynamic path of f_1' for example. This gives further evidence for the importance of initialization and diversity.

Bibliography

[Alt97] L. Altenberg. Fitness distance correlation analysis: an instructive counterexample. In Th. Bäeck, editor, *Proceedings of the 7^{th} International Conference on Genetic Algorithms*, pages 57–64. Morgan Kaufmann, 1997.

[Dav91] Y. Davidor. Epistasis variance: a viewpoint on ga-hardness. In G. J. E. Rawlins, editor, *Foundations of Genetic Algorithms*, pages 23–35. Morgan Kaufmann, 1991.

[DWP95] K. Mathias D. Whitley and L. Pyeatt. Hyperplane ranking in simple genetic alghoritms. In L. J. Eshelman, editor, *Proceedings of the 6^{th} International Conference on Genetic Algorithms*, pages 231–238. Morgan Kaufmann, 1995.

[HG95] J. Horn and D.E. Goldberg. Genetic algorithms difficulty and the modality of fitness landscapes. In L. D. Whitley and M. D. Vose, editors, *Foundations of Genetic Algorithms 3*, pages 243–269. Morgan Kaufmann, 1995.

[JF95] T. Jones and S. Forrest. Fitness distance correlation as a measure of problem difficulty for genetic algorithms. In L. J. Eshelman, editor, *Proceedings of the 6^{th} International Conference on Genetic Algorithms*, pages 184–192. Morgan Kaufmann, 1995.

[Kal98] L. Kallel. *Convergence des algorithmes génétiques: aspects spatiaux et temporels*. PhD thesis, Ecole Polytechnique – Palaiseau – France, 1998.

[Kau89] S. A. Kauffman. *Lectures in the Sciences of Complexity*, volume I of *SFI studies*, chapter Adaptation on Rugged Fitness Landscape, pages 619–712. Addison Wesley, 1989.

[KS97] L. Kallel and M. Schoenauer. Alternative random initialization in genetic algorithms. In Th. Bäeck, editor, *Proceedings of the 7^{th} International Conference on Genetic Algorithms*. Morgan Kaufmann, 1997.

[MdWS91] B. Manderick, M. de Weger, and P. Spiessens. The genetic algorithm and the structure of the fitness landscape. In R. K. Belew and L. B. Booker, editors, *Proceedings of the 4^{th} International Conference on Genetic Algorithms*, pages 143–150. Morgan Kaufmann, 1991.

[NL98] B. Naudts and L.Kallel. Comparison of summary statistics of fitnesss landscapes. To appear in Journal of IEEE transactions on Evolutionary Computation.

[QVSG98] R.J. Quick, V.J..Rayward-Smith, and G.D.Smith. Fitness distance correlation and ridge functions. In A.E. Eiben, Th. Bäeck, M. Schoenauer, and H.-P. Schwefel, editors, *Proceedings of the 4^{th} Conference on Parallel Problems Solving from Nature*, volume 1498 of *LNCS*, pages 77–86. Springer Verlag, 1998.

[RS95] N. J. Radcliffe and P. D. Surry. Fitness variance of formae and performance prediction. In L. D. Whitley and M. D. Vose, editors, *Foundations of Genetic Algorithms 3*, pages 51–72. Morgan Kaufmann, 1995.

[RVSK97] S. Rochet, G. Venturini, M. Slimane, and E.M El Kharoubi. A critical and empirical study of epistasis measures for predicting ga performances: a summary. In J.-K. Hao, E. Lutton, E. Ronald, M. Schoenauer, and D. Snyers, editors, *Artificial Evolution'97*, LNCS, pages 287–299. Springer Verlag, 1997.

[RW95] Colin Reeves and Christine Wright. An experimental design perspective on genetic alghorithms. In L. D. Whitley and M. D. Vose, editors, *Foundations of Genetic Algorithms 3*, pages 7–22. Morgan Kaufmann, 1995.