Monograph:

He, Liwen and Mort, N. (1998) Fitness Distance Correlation as a Measure of GA Performance. Research Report. ACSE Research Report 723 . Department of Automatic Control and Systems Engineering

Fitness Distance Correlation as a Measure of GA Performance

Liwen He and Neil Mort
Department of Automatic Control and Systems Engineering
University of Sheffield, Sheffield S1 3JD, UK

Research Report Number 723

Abstract: In this paper, the mathematical interpretation of correlation coefficient is reviewed to explain the conditions under which it operates. Using the work of Jones and Forrest (1995) on fitness distance correlation (FDC) as a measure of problem difficulty for genetic algorithms, a novel framework combining FDC with the Experimental Design perspective in statistics is proposed. It is shown that this method not only satisfies the mathematical condition of correlation coefficient, but also that it is closely relevant to genetic operators, such as crossover and mutation, and can therefore be used to predict the performance of genetic algorithms more accurately. Different well-known problems such as epistasis interactions, isolation or *needle-in-a-haystack*, high fitness variance, deceptiveness and multimodality, which make the GA search process difficult, are investigated. Experimental results show that this framework is an effective metric for GA performance on the fitness landscape, and offers useful guidance in constructing efficient genetic algorithms.

Keywords: Fitness Distance Correlation, Fitness Landscape, Genetic Algorithms, Experimental Design

1. Introduction

In earlier published work (Jones, 1995, Jones and Forrest, 1995), a measure of search difficulty, known as fitness distance correlation (FDC), is introduced to predict the performance of a GA on problems with known global optima. This work on heuristic and genetic algorithms suggests that the connection between fitness and distance to the global optima will have strong effect on search difficulty. The authors provide a statistical summary of many well-studied problems, and indicate that FDC can be a reasonably good predictor of GA performance. They also report some cases which FDC cannot detect, and advise the examination of a scatter plot of fitness versus distance as an alternative. They further suggest that a more accurate framework of FDC should be developed which utilizes genetic operators to define the distance from the optima.

Following their work, Altenberg (1997) constructed a counterexample to FDC. In his example, the FDC coefficient converges to a very small negative value, while the fitness and Hamming distance scatter plot displayed no discernible structure between them. In addition, he demonstrates that crossover-based fitness distance correlation may be used as a estimator of GA performance.

Our work is inspired by two observations:

1. Previous studies show that Hamming distance based FDC is too simple a statistical summary for the prediction of GA performance,
2. Jones and Forrest's FDC works well only for low dimension of the search space ($< 2^{20}$); it is ineffective for high dimensional search spaces.

In this paper, we propose a novel framework which combines FDC with the Experimental Design perspective of statistics, an approach that not only conforms to the mathematical condition for computing the correlation coefficient, but also is very relevant to genetic operators, such as crossover and mutation. In particular, our work considers the fitness versus distance scatter plot as a very important primary step since it provides richer information, and is thus effective for different dimensions of the problem search space. Unlike crossover-based FDC, this method works without actually running the genetic algorithm. It is simpler and more economical, requires less computation effort, and yet, achieves a better result.

Since the most difficult thing in using genetic algorithms is the representation issue and encoding of practical problems [Fleming, 1997], this method can assist the programmer to construct an efficient genetic algorithm for problem optimization.

The remainder of this paper is organized as follows: the mathematical background on correlation coefficient is reviewed in section 2, and the relationship between the FDC coefficient, fitness landscape and GA difficulty is discussed in section 3. Then the ideas behind our FDC framework are stated in section 4, and section 5 reveals that our FDC method is relevant to GA dynamics. The utilization of the FDC framework in different test problems is reported and analyzed in section 6; this paper is terminated by the conclusion and discussion in section 7. For completeness and to provide useful background information, some figures reported from the previous studies are contained in an Appendix.

2. Correlation Coefficient

Given that many mathematical formula are valid only under certain specific conditions, it is necessary to review the conditions under which the correlation coefficient is appropriate [Chaterfield, C, 1978]:

1. the mathematical justification for this correlation coefficient depends on two random variables having a *bivariate normal distribution*
2. only when the relationship between the two variables appears to be approximately linear, can we consider *linear correlation* analysis

As we know, most random variables arising from natural and social phenomena conform very well to the normal distribution. Thus in the case of two random variables, X and Y, each corresponding marginal distribution for these will conform to a normal distribution. Though it is not easy to prove theoretically that bivariate random variables (X, Y) will exactly conform to the bivariate normal distribution, it is convenient to assume that pairs of measurements follow a bivariate normal distribution in engineering practice. This statement is a basic assumption for the following discussion.

Linear correlation can provide a reasonable explanation for the two examples of a relationship between fitness and distance which is not detected by FDC in Jones and Forrest' s previous study. The fitness distance scatter plots of Long Path problems (Horn, Goldberg and Deb, 1994) and Liepins and Vose' s

Transform problem (1990) show that there is no linear correlation between two variables, one plot displays a zigzag shape and the other displays a X shape. (see Appendix figure 1 (*h*) and (*l*)). It is therefore unsuitable to utilize FDC as a statistical measure in these circumstances.

3. Relationship between the FDC Coefficient, Fitness Landscape and GA Difficulty

3.1. Interpretation of FDC Coefficient

The FDC coefficient is said to be *positive* if ' large' values of both variables Y (Fitness) and X (Distance) tend to occur together, and is said to be *negative* if ' large' values of one variable (Fitness) tend to occur with 'small' values of the other variable (Distance) and vice versa.

The FDC coefficient is said to be high (close to +1) if the points (x_i , y_i) cluster in a straight line and is said to be low (close to 0) if the observed points are widely scattered.

The variables are said to be uncorrelated if there is no relationship between them, and in this condition, the FDC coefficient is close to or equal to zero. Different types of correlation coefficient are illustrated in Figure 1.

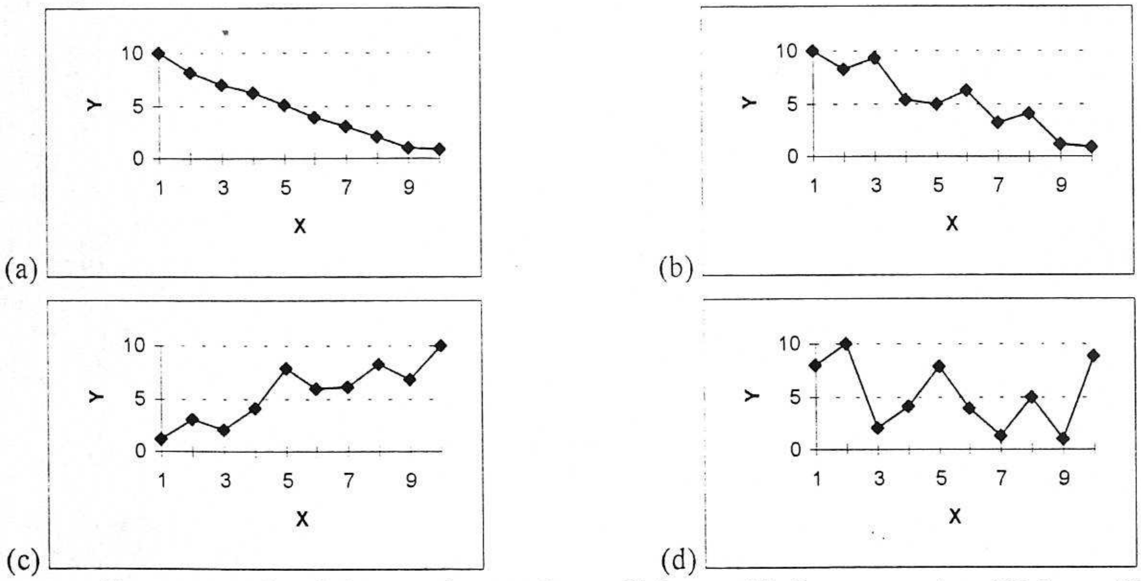

Figure 1. (a) High negative FDC coefficient (b) Low negative FDC coefficient (c) positive FDC coefficient (d) No correlation

3.2. Structure of Fitness Landscape

The fitness landscape of a genetic search is made up of directed and spatially-connected graphs. Their vertices represent the genotype and are labeled according to the fitness function, while their edges are labeled with small positive values between 0 and 1 which express the probability of moving from one vertex to another. The NK-fitness landscape defined by Kauffman (1989) is the graph where only 1-mutant neighbors are connected (for example, see Figure 2(a)), thus this fitness landscape is only pertinent to point mutation. Our fitness landscape is fully connected, that is to say, there are $\binom{Q}{2}$ edges

($Q=2^l$, l is the bitlength of genotype) in our fitness landscape (for example, see Figure 2(b)). In general, all 1-mutant, 2-mutant, 3-mutant, ..., k-mutant ($k \le l$) neighbors are relevant to the mutation operator and crossover operator (This point is discussed further in section 5). More specifically, the fitness landscape defined in this paper is a pictorial image of (l+1) dimensional space whose points are fully-connected and assigned a real fitness value. The edge are labeled with the value $p_{ij} \in [0, 1]$ (for $i \le l$ and $j \le l$) which represents the probability of moving from one vertex i to another vertex j. For example, if the genetic operators are universal crossover and point mutation, the maximum Hamming distance between parent and offspring is ($l/2+1$), so the probability of moving between any two points whose Hamming distance is greater than ($l/2+1$) is equal to zero.

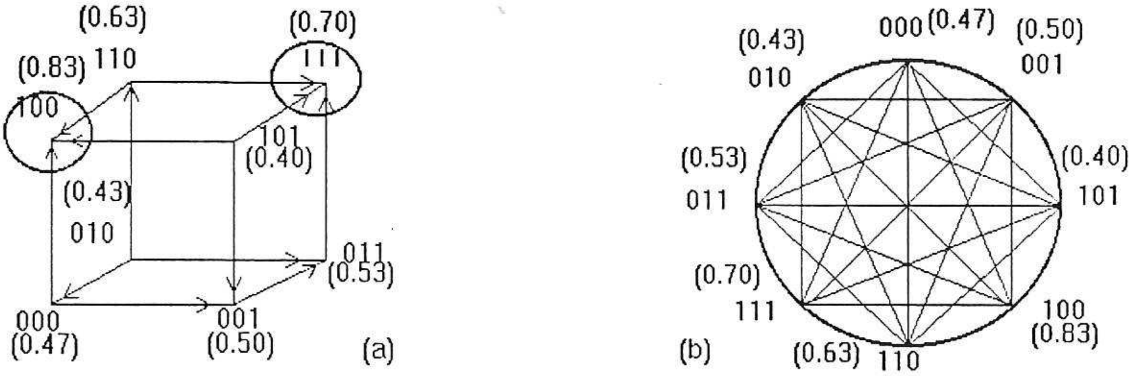

Figure 2. Fitness Landscape with dimension 2^3 of space generated by:(a) point mutation (b) crossover plus point mutation

From the viewpoint of genetic search, a fitness landscape is regarded as mountainous: that is, those vertices with best fitness value are named as ' peaks' or global optima, and a population would be expected to be moved by genetic operators such as selection, mutation and recombination to those peaks since genetic algorithms search for optima. The point with fitness value greater than all its neighboring points is termed a local optimum, for example, the vertex 111 in figure 2(a).

An important measure of fitness landscape is its *smoothness* and *ruggedness*, or in more precise terms, the ' correlation structure' of the fitness landscape. A *smooth* or *correlated* landscape means neighboring points in the genotype space have nearly the same fitness values. The fitness value of one point will provide relatively accurate information about the fitness values of neighboring points. On the other hand, a *rugged* or *uncorrelated* landscape indicates very large fitness differentials between neighboring points, and the fitness value of one point carries no information about the expected fitness value of neighboring points. However, different landscapes have different definitions of the neighboring relationship. The neighboring points in the NK-fitness landscape defined by Kauffman (1989) are 1-mutant. The neighboring points in our approach are 1-mutant, 2-mutant, ... k-mutant ($k \le l$).

3.3. Measurement of GA Difficulty

In the fundamental theory of GAs, a number of factors have been addressed from different viewpoints which prevent GA search achieving global optimization, such as spurious correlation or hitch-hiking, (Scaffer, et al, 1991; Forrest & Mitchell, 1993; Mitchell, 1996), epistasis interaction (Davidor, 1991; Rochet, 1997), isolation or *needle-in-a-haystack* (Horn and Goldberg, 1995), high fitness variance(Grefenstette, 1993), deceptiveness, (Goldberg, 1987; Whitley, 1991), and multimodality (Ohkura and Ueda, 1995; He & Mort, 1998a).

Unfortunately, there is no accurate mathematical model to measure the problem difficulty for genetic algorithms, as each of the above elements represent just a part of the total picture. Some of them overlap with each other. For example, in the NK-fitness landscape (Kauffman, 1989), increasing the degree of epistatis interaction, K, increases the ruggedness of fitness landscape, but also, increases the multimodality--the number of local optima become larger, and the fitness variance also increases. Most multimodal optimization problems have some kind of deceptiveness property, and can be regarded as the most deceptive problem. (Ohkura and Ueda, 1995; He and Mort, 1998a).

Other promising approaches to the GA difficulty measurement problem are computing the autocorrelation of fitnesses obtained from a random walk on the landscape (Weinberger, 1990), and the correlation length of landscape, (Manderick, *et al*, 1991). More recently, Hordijk (1995) introduces the autocorrelation of a time series (known as the Box-Jenkins approach) of fitness values generated by the random walk on the landscape.

These metrics, although effective in investigating the GA difficulty, are not easily utilized in practice as most of them are more complex than using genetic algorithms themselves. The objective of this paper is to provide a simple and economical evaluation methodology to measure GA performance, and then use it rapidly construct genetic algorithms for engineering optimization problem solving.

One statistical metric which characterizes the correlation structure of fitness landscape is fitness distance correlation coefficient. Preliminary results suggest that it can provide a reliable indication of problem difficulty for genetic algorithms. For a maximization problem, the closer to zero the FDC coefficient is, the more rugged the fitness landscape is, and so the more difficult is the genetic algorithms search for the optimum.

4. New FDC Framework with Experimental Design

Genetic search can be regarded as a process of exploitation on directed and spatial graphs whose vertices are assigned by the fitness function. These graphs are the fitness landscape. FDC is a simple method to measure the relationship between fitness and distance to the global optimum, where the distance is approximated by the Hamming distance under the actual genetic operators. Given a set $\boldsymbol{F} = \{f_1, f_2, ..., f_n\}$ of n individual fitnesses and a corresponding set $\boldsymbol{D} = \{d_1, d_2, ... , d_n\}$ of the n distance to the nearest global maximum, the fitness distance correlation coefficient r is calculated as:

$$r = c_{FD} / (s_F \, s_D)$$

where

$$c_{FD} = \sum_{i=1}^{n} (f_i - f_m)(d_i - d_m)$$

is the covariance of F and D, and s_F, s_D, f_m and d_m are the standard deviations and means of $\boldsymbol{F}$ and $\boldsymbol{D}$ respectively. 2^{12} or 4000 randomly points are sampled from the fitness landscape as the candidate individuals in the work of Jones and Forrest (1995).

Experimental design is a traditional statistical methodology. It extracts information from a number of carefully selected, representative, but, randomly-produced data sets, and attempts to account for the observed phenomena explicitly. This perspective requires substantial human interaction and interpretation in the light of the particular problem being examined. (Chaterfield, 1978; Reeves and Wright, 1995)

A new FDC framework with experimental design is proposed here. It is inspired by the phenomenon that Jones and Forrest' s FDC cannot work well in a large search space, such as the 64 bit Royal Road function (Mitchell, 1996). Figure 3 shows the fitness distance scatter plot of the Royal Road function

RR1 according to the Jones and Forrest FDC method. From the viewpoint of Hamming distance axis, it is clear that the randomly generated 4000 points are symmetrical about the line HD=$l/2$=32 on the Hamming Distance axis, and situated within the interval [20 45]. That is to say, their FDC just contains the information about part of the entire Hamming distance region in fitness distance scatter plot. A large amount of information about the fitness values close to or far away from the global optimum as measured by Hamming distance has been lost. This information is very critical to the FDC framework, and without it, there is a high probability of drawing invalid conclusions. According to Figure 2.(d) provided by their study on the 128 bit Royal Road function (see **Appendix 2**), the 4000 random points are also limited within the interval [40 88] on the Hamming distance axis.

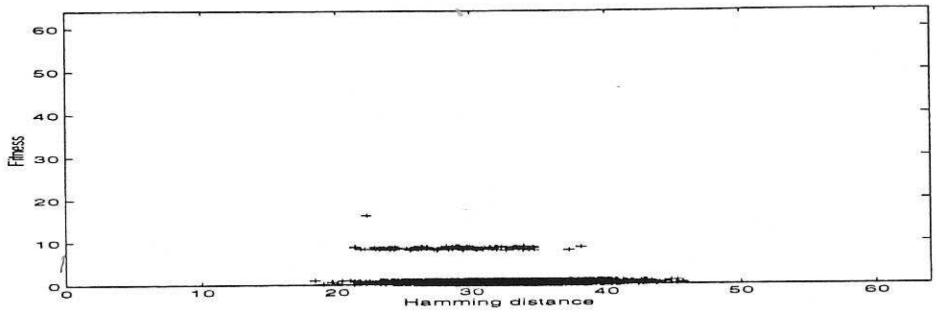

Figure 3. fitness distance scatter plot of 64 bits Royal Road function RR1 (Jones and Forrest FDC method)

On the basis of the above discussion, a novel FDC framework is proposed and stated as follows (using a search space greater than or equal to 2^{63} as an example):

1. generate 4000 random points, marking their intervals along the Hamming distance axis as [d_1 d_2];
2. generate 3000 points in the interval of [1 d_1] and [d_2 l] respectively along the Hamming distance axis
3. calculate the corresponding FDC coefficient at these three region as r_L, r_M, and r_H.
4. compare the values of r_L, r_M, and r_H for different 'benchmark' problems

Special attention should be paid to step 2, where the experimental design perspective is implemented. Our target is to generate evenly distributed points within intervals [1 d_1] and [d_2 l], and only the mutation operator is used here for simplicity.

When the mutation operator is applied to the global optimum a number of times, the simulation results show that the frequency distribution of the offspring is approximately a discrete normal distribution with mean value equal to ($mu*l$) measured along the Hamming distance (to the global optimum) axis. Here, mu is the mutation rate, and l is the binary length of the global optimum. Therefore a number of carefully selected values of mutation rate can be used to produce approximately evenly distributed points within the interval of [1 d_1] and [d_2 l]. These control mutation rate values are named as x_1, x_2, ..., x_p and y_1, y_2, ..., y_p respectively.

Jones (1995) suggested that Hamming distance is strongly related to the mutation operator in traditional genetic algorithms which use a binary string representation. The number of times that a mutation operator must be utilized to transform a given string to the global optimum is monotonic with Hamming distance. But the crossover operator is regarded as a more important genetic operator; mutation is just a safety strategy to prevent premature convergence in classical genetic algorithms. This suggests again that the previous FDC framework is a too simplistic for the GA performance assessment.

In the next section, we will show that our FDC framework is closely related to the crossover operator, and the dynamic property of FDC is clarified.

5. Dynamic Property of the Novel FDC Framework

Jones and Forrest (1995) recommend that a stronger predictor of GA optimization performance would be FDC analysis using distance measures based on the genetic operators themselves. Altenberg (1997) constructs two kinds of crossover-based fitness distance correlation analysis methods. The relevant fitness distance scatter plots can be found at Appendix 2. This approach works well for some applications.
Now, the relationship between our novel FDC framework and the GA dynamic property is made clear.

First, the role of genetic operators such as selection, crossover, and mutation in the evolutionary procedure is reviewed. An initial population is generated randomly, pairs of bitstrings are chosen as the parents to produce the offspring through crossover, then mutation is implemented to every single bitstring. Though there exist many different crossover methods such as single point crossover, double point crossover, multi-point crossover, universal crossover, etc, and the resulting offspring are dependent upon both the particular crossover method used and the composition of the population, any offspring in the next generation can be regarded as flipping some specific bits of the parents in the previous population. This process can be simulated equivalently by the inverse operation: the mutation operator is applied to the global optimum with different values of control mutation rate. This operation is indicated in the section 4. Therefore, our fitness distance scatter plot can be considered as a statistical summary of the random sampling points from an evolutionary procedure of applying genetic operators onto the entire population.

6. FDC Experimental Results and Analysis

As indicated in the previous section, a number of well-studied problems which represent GA difficulty have been selected as candidates for our research. These problems include epistasis interaction, isolation or *needle-in-a-haystack* (NIAH), high fitness variance, deceptiveness and multimodality. The test functions are summarized in the Appendix. Since the spurious correlation or hitch-hiking problem comes from the characteristics of the genetic algorithms themselves, the FDC framework cannot detect it.

However, before any further investigation, it must be emphasized that any comparison between problem difficulty for GA optimization should be based on these being approximately equal dimension in each problem search space, since it is well-known that problem difficulty is proportional to the input size of the GA search space.

Therefore, our test problems are the 64 bit Royal Road function RR1 and RR2 (Forrest and Mitchell, 1993), 16 copies of 4 bits Whitley' s (1991) F2 fully deceptiveness function (total 64 bits), NIAH1 function (Jones and Forrest, 1995), NIAH2 function (Grefenstette, 1993), FDC counterexample (Altenberg, 1997), maximum multimodality (Horn and Goldberg, 1995) and the combinatorial multimodal optimization problem (63 bits) (He and Mort, 1998a, 1998b) which are of the approximately same dimension of search space. According to the experimental design perspective indicated in section 4, the FDC framework is computed in the following procedures:

1. Randomly generate 4000 points which is situated within the interval of [20 45] on the Hamming distance axis. The fitness distance correlation coefficient r_M in this interval is calculated, making r_M the FDC coefficient in the Jones and Forrest approach.

2. Mutation of the global optimum is used to produce approximately evenly distributed points within the interval of [1 20] and [20 l], and the control mutation rate values are carefully selected as 1.5/l, 4.0/l, 6.5/l, 9.0/l, 10.5/l, 13.5/l, 16.5/l, and 47.5/l , 50.5/l,53.5/l, 56.5/l, 59.5/l, 62.5/l. The respective FDC coefficients are calculated as r_L and r_H.

The fitness distance scatter plots are summarized in Figure 4. (a) to (d), and the FDC coefficients are displayed in table 1.

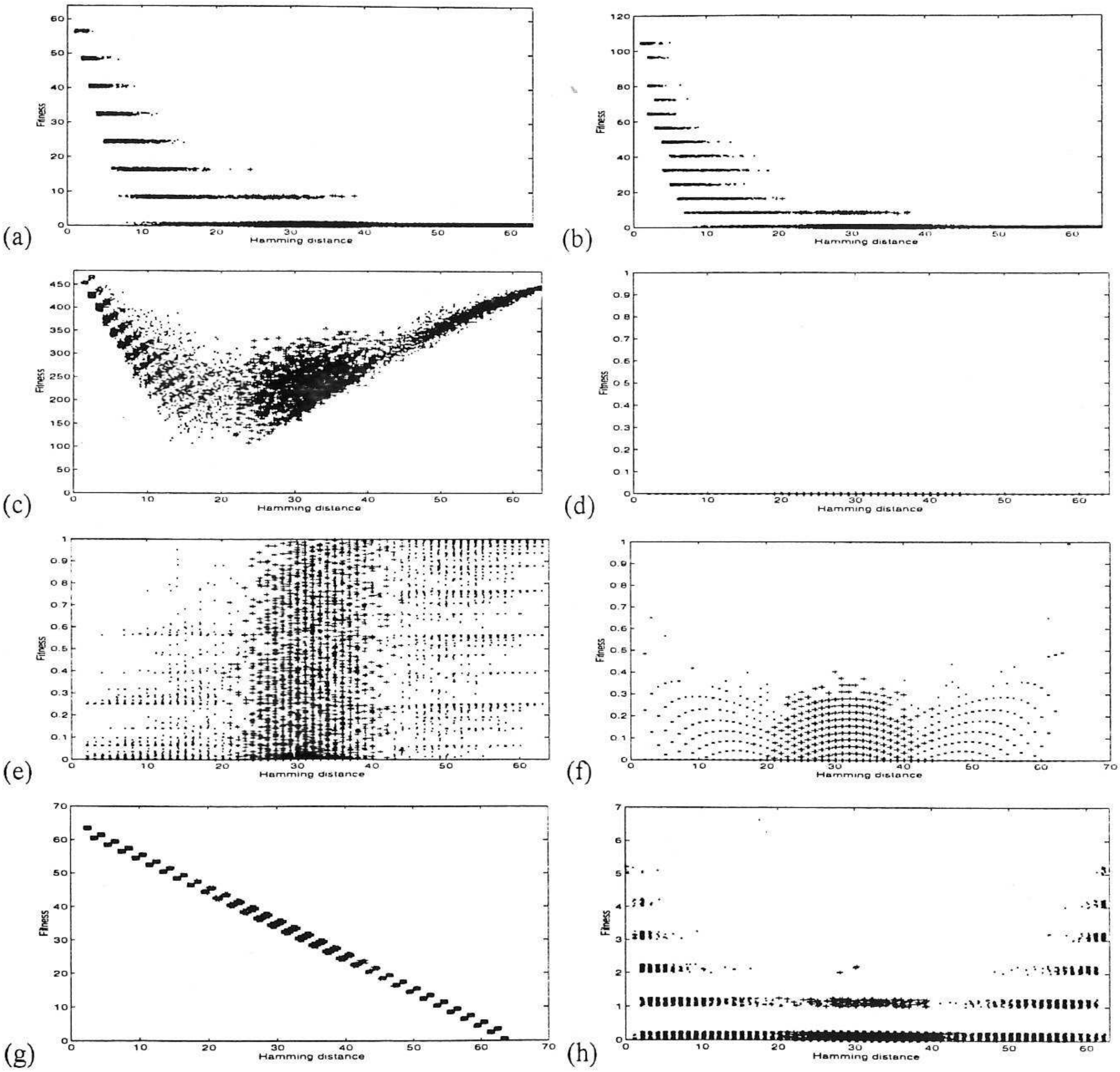

Figure 4. Fitness distance scatter plot

(a) (b) 64 bits Royal Road function RR1 and RR2;
(c) 16 copies of 4 bits Whitley's F2 fully deceptiveness function (d) NIAH1 function
(e) NIAH2 function (f) FDC counterexample
(g) maximum multimodality (63 bit) (h) combinatorial multimodal optimization problem (63 bits)

(Note: Since the axis values of the points in the fitness distance scatter plot are integer, in order to display the distribution of points clearly, visualization programming is introduced here by adding a small random value within the range 0 $\rightarrow$ 0.25 to both the fitness and Hamming distance values.)

	r_L	r_M	r_H
RR1	-0.8924 (0.0017)	-0.1766 (0.0160)	None
RR2	-0.8368 (0.0031)	-0.1767 (0.0160)	None
Deceptive(Whitley F2)	-0.8799 (0.0033)	0.5112 (0.0139)	0.9718 (0.0009)
NIAH1	None	None	None
NIAH2	0.3355 (0.0156)	0.2114 (0.0167)	0.4493 (0.0117)
FDC Counterexample	0.0448 (0.0160)	-0.0024 (0.0138)	0.2541 (0.0155)
Maximum Modality	-0.9848 (0.0002)	-0.9695 (0.0006)	-0.9856 (0.0003)
Combinatorial Multimodality	-0.5951 (0.0067)	-0.0022 (0.0168)	0.7143 (0.0060)

Table 1. FDC coefficient of all Test Functions

The above FDC coefficient values are the mean values of ten computations of FDC on the different test functions with the variance contained in parenthesis. The variance in our novel FDC framework is very small, which means that the fitness landscape for the ' random walks' is *statistically isotropic*. A statistically isotropic landscape is one in which the statistics of the sequence of fitness values are the same regardless of the starting point selected. More formally, these fitness landscapes are those in which fitness values assigned to the configurations also form a stationary random process for the assumed joint distribution of fitnesses (Weinberger, 1990). That is to say, all the test problems satisfy the condition, so the random walks on these fitness landscapes can be regarded as stationary random processes.

The fitness scatter plots in figure 4 shows that they can display approximate linearity at the three different intervals of Hamming distance, therefore, the utilization of FDC coefficient is achieved with some confidence.

The interpretation of FDC coefficient as a measure of GA difficulty is reviewed here. According to our previous discussion, r_M represents the tendency of a population to move to the target optimum (r_M negative) or in the reverse direction (r_M positive) when a random initial population is generated. r_L and r_H. are closely connected to the GA dynamic property indicated by the use of genetic operators such as crossover and mutation. They predict the trend of a population' s movement which is guided by the fitness function after a few evolutionary generations.

Without loss of generality, the problems considered are always maximization. The global optimum is always at the point of (0, FV_{max}) in the fitness distance scatter plot (FV_{max} is the fitness value of the global optimum). Thus, linking with the previous discussion in section 3, the following rules are generated:

1. high negative FDC coefficient (close to -1) represents a smooth fitness landscape, and it is easy for genetic algorithms to search for the global optimum
2. low negative FDC coefficient means a rugged fitness landscape, and it is relatively difficult for genetic

algorithms to search for the global optimum

3. population evolving on the fitness landscape with positive FDC coefficient will drift to a local optimum or another global optimum far away from the target global optimum. Thus a positive FDC coefficient represents GA deceptiveness, the higher the positive value (closer to +1), the more deceptive the problem.

Now the results of the FDC coefficient computations for the test problems in table 1 are analyzed by considering also their respective fitness distance scatter plots in figure 4.

1. The FDC coefficients in these two function are consistent with the GA simulation results reported in Forrest and Mitchell (1993). (see Appendix 1). All the values of r_L and r_M in the RR1 function and RR2 function are negative, and r_H is ' None' , so they represent non-deceptive problems. From the analysis of Rochet (1997), the epistasis interaction in the RR2 is greater than that of RR1. It is seen that the r_L value provides confirmation of this since r_L in RR1 function is smaller than that of RR2 function, so our FDC can accurately predict the epistasis interaction in genetic algorithms applications. Note also that r_M is the measure used in the Jone and Forrest approach, and produces values that are virtually equal for RR1 and RR2 functions, so it cannot be used satisfactorily an indicator of GA search difficulty for Royal Road functions.

2. In the results presented, the value of r_H for Whitlety's F2 fully deceptive function is high positive, and greater than the absolute value of r_L. That is to say, the fitness landscape around the local optimum is smoother than that around the global optimum, therefore the population is moved away from the global optimum to a local optimum. Our FDC framework therefore accurately predicts the deceptiveness problem from the value of r_H.

3. 'Needle-in-a-Haystack' is regarded as a GA-Hard problem traditionally because of its fitness isolation and high fitness variance. Two types of NIAH problem are considered in this paper. The fitness value of the NIAH1 function is zero everywhere except for one point. Our FDC framework does not include the optimal point because the mutation operator is applied to the global optimum, so all these FDC coefficients are non-existent (None), and not equal to the zero as indicated by Jones and Forrest (1995). This result shows that since FDC coefficient does not exist, the GA used in this function is no more than a random search method since no reference to fitness variance is made.
   The values of FDC coefficients of the NIAH2 function are all positive, so this function is just a GA deceptive problem, and therefore is difficult for GA search. This result confirms the report in Grefenstette (1993) that if the optimum is not in the initial population, it will never be found by a GA search. From our FDC analysis, it is seen that the secret of NIAH2 is the deceptiveness problem. Isolation and high fitness variance in the global optimum is not a necessary condition for GA search difficulty, since the GA search is irrelevant to the fitness value of the global optimum.

4. The values of r_L and r_H in Altenberg's (1997) FDC counterexample are positive, which means that this function represents the deceptiveness problem in general, and given that r_L is close to zero, the fitness landscape around the global optimum is extremely rugged, so this function is very difficult for genetic algorithm search. This analysis is in accordance with the simulation result in Altenberg' s paper: even for the 50 bit function, the mean number of function evaluations in finding the optimum at 41% success rate is 10^6; it is reasonable to assume that the number of function evaluations of the 64 bit function would be greater than 10^6. Compared with the mean number of function evaluations to find the optimum by GA at 100% success rate on the 64 bit Royal Road function RR1 (Mitchell, 1996, P130) given as 61334, this counterexample is very difficult for a GA search.

5. All FDC coefficients of Maximum Modality are high negative (close to -1), which represents a

very smooth fitness landscape and it is very easy to search for the global optimum by GA. This analysis is fully in keeping with the simulation results in which 39 out of 40 runs of GA for the bitlength 49 function succeed in converging to a global optimum in Horn and Goldberg (1995). Also, it should be made clear that massive multimodality is all the local optima whose fitnesses is smaller than the single global optimum, so non deceptiveness is represented in this function which is therefore easy for GA search.

6. The value of r_H in the Combinatorial Multimodal Optimization Problem is positive, though the value of r_L is negative. Moreover, the absolute value of r_H is greater than that of r_L , so for any definite global optimum, it is not easy to search by GA as deceptiveness is apparent. This analysis is in accordance with the GA simulation result on this function (He and Mort, 1997). Also, it should be pointed out that while the Maximum Modality function has only one global optimum, the Combinatorial Multimomdal Optimization Problem (CMOP) has a number of global optima and a large number of local optima, so CMOP can be regarded as a most deceptive problem (Ohkura and Ueda,1995; He and Mort, 1998a,b).

7. Conclusion and Discussion

On the basis of the pioneering work of Jones and Forrest (1995) on FDC as a measure of problem difficulty for genetic algorithms, a novel framework combining FDC with the Experimental Design perspective in statistics is proposed. This approach not only meets the mathematical condition of correlation coefficient, but also is closely relevant to genetic operators, such as crossover and mutation. A number of different well-known test functions with particular GA difficulty factors including epistasis interaction, isolation, high fitness variance, deceptiveness and multimodality are investigated. Numerical results demonstrate that this framework is an effective statistical summary for GA performance on the respective fitness landscape.

Though this work uses the 63 or 64 bits test functions as examples, other problems with different dimension of search space can be addressed in a similar way. Also, it should be recognized that the FDC coefficient is merely a relative value used to compare predicted GA performance on the different fitness landscapes with approximately the same dimension of search space. We do not recommend a special value of FDC coefficient to be a standard for GA search difficulty, only the relative performance is measured using our FDC methodology.

Since the fitness landscape is related to different encoding or representation issues and the assignment of fitness functions for GA applications, this framework provides simple and effective guidance to construct genetic algorithms for practical engineering problems.

References

Alternberg, L. (1997): Fitness Distance Correlation Analysis: An Instructive Counterexample. *Proceedings of 7th International Conference on Genetic Algorithms*. Michigan State University. Thomas Back (Ed.). Morgan Kaufmann. 57-64.

Box, G.E.P and Jenkins, G.M. (1970): *Times Series Analysis, Forecasting and Control*. Holdern Day.

Chaterfield, C. (1978): Statistics for technology: a course in applied statistics. John Wiley and Sons, New York.

Davidor, Y. (1991): Epistasis Variance: A view point on GA-hardness. *Foundation of Genetic*

Algorithms, Vol 1. Rawlins, G.J.E (Ed). San Mateo, CA: Morgan Kaufmann. 23-35.
Fleming, P.J (1997): Personal communication.
Forrest, S., and Mitchell, M. (1993): Relative Building-Blocks Fitness and Building-Blocks Hypothesis. *Foundation of Genetic Algorithms, Vol 2.* Vail Colorado: Morgan Kaufmann. 109-126.
Goldberg, D.E. (1987): Simple Genetic Algorithms and the Minimal, Deceptive Problem. Chapter 6 of *Genetic Algorithms and Simulated Annealing.* Davis, L.(Ed). London: Pitman. 74-88.
Grefenstette, J. (1993): Deception Considered Harmful. *Foundations of Genetic Algorithms. Vol.2.* Vail Colorado: Morgan Kaufmann. P75-91.
He, L and Mort, N. (1998a): Intelligent Genetic Algorithms in Evolutionary Computation. Part I. Biological Foundation; Part II. Application to Combinatorial Multimodal Optimization Problem. ACSE Department Research Report 689, 690. University of Sheffield. Submitted to *Evolutionary Computation.*
He, L and Mort, N. (1998b): Modelling and Optimal Scheduling of Distributed Computer Control Manufacturing System. Submitted to IFAC World Congress, 1999, Beijing.
Hordijk, W. (1995): A Measure of Landscapes. Santa Fe Institute (SFI) Working Papers 95-05-049.
Horn, J., Goldberg, D.E, and Deb, K. (1994): Long Path Problems. *Parallel Problems Solving From Nature - PPSFN III, volume 866 of Lecture Notes in Computer Science.* Davidor, Y et al (Eds). Berline: Springer-Verlag. 149-158.
Horn, J and Goldberg, D. E. (1995): Genetic Algorithm Difficulty and the Modality Fitness Landscape. *Foundations of Genetic Algorithms. Vol. 3.* Whitley, D. and Vose, M.D (Eds). Morgan Kaufmann. P243-267.
Jones, T.(1995): *Evolutionary Algorithms, Fitness Landscapes and Search.* PhD thesis. University of New Mexico, March 1995.
Jones, T. and Forrest, S, (1995): Fitness Distance Correlation as a Measure of Problem Difficulty for Genetic Algorithms. *Proceedings of 6th International Conference on Genetic Algorithms.* University of Pittsburgh. Larry, J. Eshelman (Ed.).184-192.
Kauffman, S.A (1989): Adaptation on Rugged Fitness Landscape. *Lectures in the Science of Complexity, SFI Studies in the Science of Complexity.* D. Stein (Ed.), Addiso-Wesley, Longman.
Liepins, G., and Vose, M.D. (1990): Representational Issues in Genetic Optimization. *Journal of Experimental and Theoretical Artificial Intelligence*, **2**, 101-115.
Manderick, B., Weger, M. De, and Spiessens, P. (1991): The Genetic Algorithms and the Structure of the Fitness Landscape. *Proceedings of 44th International Conference on Genetic Algorithms.* San Mateo, CA: Morgan Kaufmann. Belew, R.K and Booker, L.B.(Eds). 143-150.
Mitchell, M. (1996): *An Introduction to Genetic Algorithms.* The MIT press.
Ohkura, K and Ueda, K. (1995): A Genetic Algorithms with Neutral Mutations for Massively Multimodal Function Optimization. *IEEE International Conference on Evolutionary Computation*, Australia, PP.361-366.
Reeves, C.R. and Wright, C.C. (1995): Genetic Algorithms and Statistical Methods: A Comparison. *IEE Conference on Genetic Algorithms in Engineering Systems.* University of Sheffield. 137-140.
Rochet, S. (1997) : Epistasis in Genetic Algorithms Revisited. *Information Science* **102**, 133-155.
Scaffer, J.D, Eshelman, L.J., and Offutt, D. (1991): Spurious Correlations and Premature Convergence in Genetic Algorithms. *Foundation of Genetic Algorithms, Vol 1.* Rawlins, G.J.E (Ed). San Mateo, CA: Morgan Kaufmann. 102-112.
Weinberger, E. (1990): Correlated and Uncorrelated Fitness Landscapes and How to Tell the Difference. *Biological Cybernetics.* **63**. 325-336.
Whitley, L.D. (1991): Fundamental Principles of Deception in Genetic Search. *Foundation of*

Genetic Algorithms, Vol 1. Rawlins, G.J.E (Ed). San Mateo, CA: Morgan Kaufmann. 221-241.

Appendix

1. Test Functions and their reported results

(1) Royal Road functions RR1 an RR2 are shown below.

Royal Road function RR1

s_1 =11111111**; c_1= 8
s_2 =********11111111**; c_2= 8
s_3 =****************11111111**; c_3= 8
s_4 =************************11111111********************************; c_4= 8
s_5 =********************************11111111************************; c_5= 8
s_6 =**11111111****************; c_6= 8
s_7 =**11111111********; c_7= 8
s_8 =**11111111; c_8= 8
s_{opt} =1111111111111111111111111111111111111111111111111111111111111111; c_{opt}= 64

(2) Royal Road function RR2

s_1 =11111111**; c_1 = 8
s_2 =********11111111**; c_2 = 8
s_3 =****************11111111**; c_3 = 8
s_4 =************************11111111********************************; c_4 = 8
s_5 =********************************11111111************************; c_5 = 8
s_6 =**11111111****************; c_6 = 8
s_7 =**11111111********; c_7 = 8
s_8 =**11111111; c_8 = 8
s_9 =1111111111111111**; c_9= 16
s_{10} =****************1111111111111111********************************; c_{10}= 16
s_{11} =********************************1111111111111111****************; c_{11}= 16
s_{12} =**1111111111111111; c_{12}= 16
s_{13} =11111111111111111111111111111111********************************; c_{13}= 32
s_{14} =********************************11111111111111111111111111111111; c_{14}= 32

s_{opt} =1111111111111111111111111111111111111111111111111111111111111111; c_{opt}= 192

In RR1 function, the given order-8 schema is a building block, the fitness value RR1(s) is the sum of the coefficient c_s corresponding to each given schema of which bit string *s* is an instance. For example, RR1(1111111100...0)=8, RR1(1111111111111111100...00)=16, and RR1(s_{opt})=64.

In RR1 function, some certain intermediate stepping stones with higher fitness contribution are added based on RR1 function, RR2 is calculated in the same way as RR1: the fitness of a bit string *s* is the sum of the coefficient corresponding to the each schema (s_1-s_{14}) of which it is an instance. For example, RR1(1111111100...0011111111)=16, since this string is an instance of both schema s_1 and s_8, but RR1(1111111111111111100...00)=32 since this string is an instance of s_1, s_2 and s_9. Also the optimal fitness value is 192, since it is an instance of all schema in the list. The reported GA simulation result is:

	Function Evaluations to Optimum	
500 runs	RR1	RR2
Mean	62099	74563

(3) Whitley's 4 bit fully deceptive function F2

This function is such that all low order hyperplane (order 4 schema) information leads the search away from global optimum 1111 to local optimum (deceptive attractor) 0000.

f(1111)=30	f(0100)=22	f(0110)=14	f(1110)=6
f(0000)=28	f(1000)=20	f(1001)=12	f(1101)=4

f(0001)=26	f(0011)=18	f(1010)=10	f(1011)=2
f(0010)=24	f(0101)=16	f(1100)=8	f(0111)=0

(4) Needle in a Haystack function NIAH1:

The fitness value is zero everywhere except for one point.

(5) Needle in a Haystack function NIAH2:

Consider an L-bit space representing the interval [0 1] in binary encoding. Let the fitness function f be defined as:

$$f(x) = 2^{(L+1)} \quad \text{if } x = 0$$
$$f(x) = x^2 \quad \text{otherwise}$$

It is almost impossible for a GA to find the optimum if the initial population does not include the optimum in both the above NIAH functions.

(6) FDC counterexample: consider the following fitness function of bit string x:

$$F(x) = \max\,[1 - D(x)\,L\,/\,\{2\,[L-H(x)]H(x)\}\,,\ 0]$$

where, $D(x) = \sum_i^{L-1} |x_i - x_{i+1}| \in [0, L-1]$;

$H(x)$ is the Hamming distance to the global optimum $\mathbf{0}$, $\mathbf{0}$ is the all 0s bitstring which has the maximum fitness value $F(0) = 1 - 1/[2\,(L-1)]$. The reported GA experimental results are:

Bit length	Mean Evaluations	Success
24	56.3K	100%
32	164K	93%
50	10^6	41%

(7) In order to illustrate that massive multimodality by itself does not imply difficulty for GA search, a maximally multimodal function which has the maximum number of possible ' attractors' but is easy for GA search may be constructed as:

$$f_{mm.easy}(s) = u(s) + 2 * f_{mm}(s)$$

where, $f_{mm}(s) = 1$ if odd $(u(s))$; 0 otherwise

and $u(s)$ is the unitation of a string s which is equal to the number of 1s in s. For example, $u(01011)$ =3, is odd. The GA simulation results is:

Problem Size	No. of Trials (of 40) converging to optimum
l =29	40
l =39	39
l =49	39

(8) Combinatorial Multimodal Optimization Problem (CMOP):

CMOP comes from a practical manufacturing scheduling problem. A number of integers is characterized by the vector W= {176 380 216 688 144 497 153 12 714 231 310 170 6 660 50 114 282 12 454 128 266} . A great integer is given as P=811. Some of these integers will be sampled without replacement to form the *combinations* whose sum is close to P but no more than P. The fitness value is assigned to be the number of the combinations subject to the constraint condition:

$$\sum_{i=1}^{m} C_i = P\,m - \text{sum}(W)$$

where C_i is i-th combination. m is the number of combinations. This is a typical combinatorial optimization problem with at least two optima as shown below:

$s_1=$

497	12	380	144	176
310	231	153	660	216
0	114	12	6	282
0	454	266	0	128
807	*811*	*811*	*810*	*802*

sum(C_i) *807* *811* *811* *810* *802*

$s_2=$

216	12	380	144	176
310	231	153	6 -	497
282	114	12	660	128
0	454	266	0	0

sum(C_i) *808* *811* *811* *810* *801*

So the optimal fitness value m=5, just 5 combinations can be found from these integers which satisfies the constraint condition. This combinatorial multimodal optimization problem has numerous global optima. The results of 50 trials of a simple GA using 30,000 function evaluations are reported as follows:

Final fitness value until maximum generation	5	4	3
Number of fitness value found in 50 trials	18	26	6

2. Fitness Distance Scatter Plots reproduced from Jones and Forrest (1995), (Figure 2)

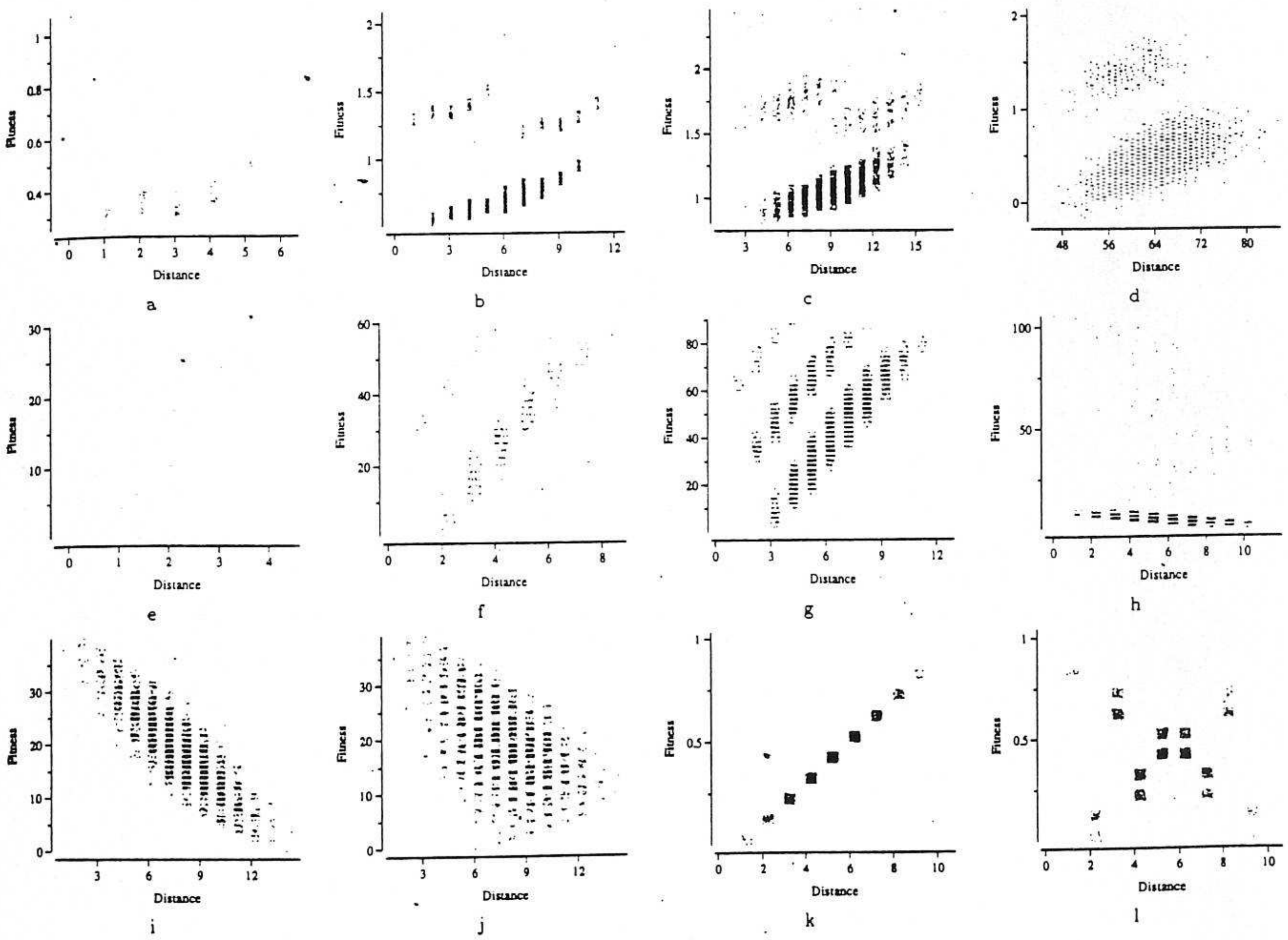

Figure 2: A sample of fitness distance scatter plots. Function sources are given in Table 1. FDC values for functions on more than 12 bits are computed from a random sample of 4000 points. The functions are as follows: (a-c) one, two and three copies of Deb & Goldberg's fully deceptive 6-bit problem ($r = 0.30$). Notice the additive effect. (d) Holland's royal road on 128 bits ($b = 8$, $k = 4$ and $g = 0$), ($r = 0.27$). (e-g) one, two and three copies of Whitley's F2, a fully deceptive 4-bit problem ($r = 0.51$). (h) Horn, Goldberg & Deb's long path problem with 11 bits ($r = -0.12$). Notice the path. (i,j) De Jong's F3 binary and Gray coded with 15 bits as a maximization problem ($r = -0.86$ and -0.57). (k) Liepins and Vose's fully deceptive problem on 10 bits ($r = 0.98$) and (l) their transformed problem ($r = -0.02$). Correlation cannot detect the X.

3. FDC counterexample reproduced from Altenberg (1997)

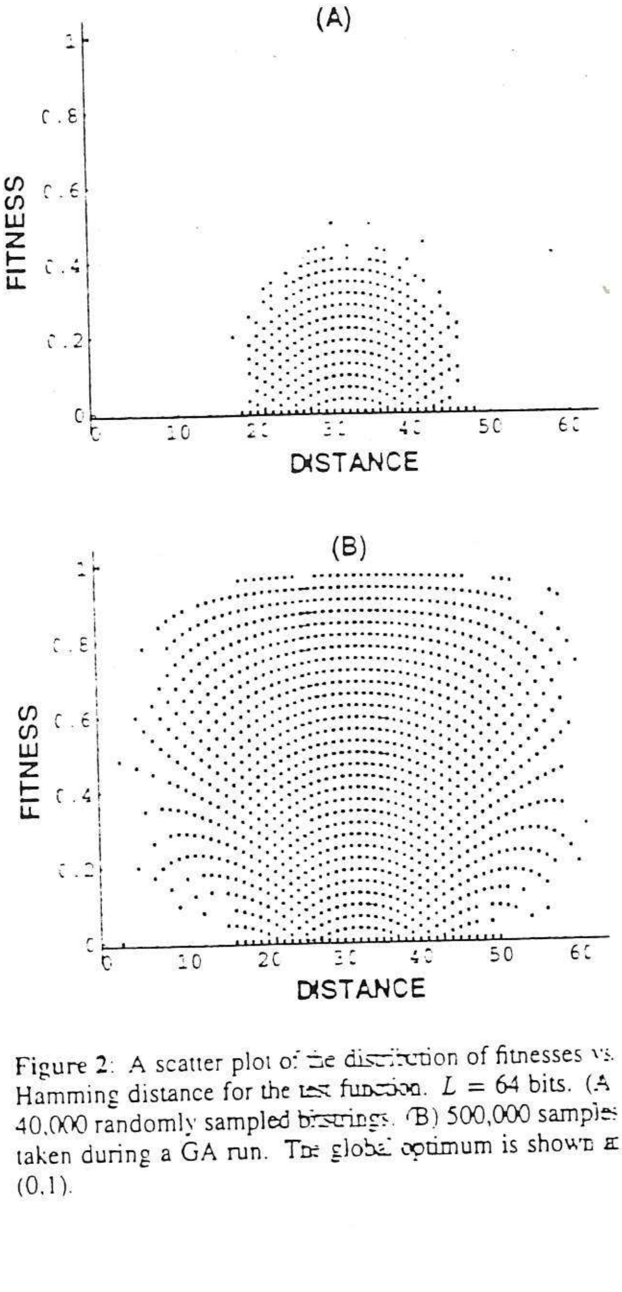

Figure 2: A scatter plot of the distribution of fitnesses vs. Hamming distance for the test function. $L = 64$ bits. (A) 40,000 randomly sampled bitstrings. (B) 500,000 samples taken during a GA run. The global optimum is shown at (0,1).

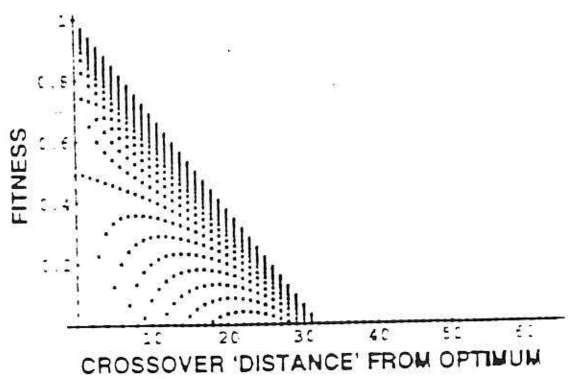

Figure 4: A plot of the test fitness function versus crossover distance—the number of discontinuities between 0s and 1s in the bitstring. The bitstring length is $L = 64$. Values for the entire search space are plotted.