[Next] [Up] [Previous]

Next: [Evolutionary Algorithm](#) **Up:** [CIXL2: A Crossover Operator](#) **Previous:** [Crossover Dynamics](#)

Benchmark Problems

In the field of evolutionary computation, it is common to compare different algorithms using a large test set, especially when the test involves function optimization [GW93]. However, the effectiveness of an algorithm against another algorithm cannot be measured by the number of problems that it solves better. The ``*no free lunch*'' theorem [WM95] shows that, if we compare two searching algorithms with all possible functions, the performance of any two algorithms will be , on average, the same . As a result, attempting to design a perfect test set where all the functions are present in order to determine whether an algorithm is better than another for every function, is a fruitless task.

That is the reason why, when an algorithm is evaluated, we must look for the kind of problems where its performance is good, in order to characterize the type of problems for which the algorithm is suitable. In this way, we have made a previous study of the functions to be optimized for constructing a test set with fewer functions and a better selection [WMRD95,Sal96]. This allows us to obtain conclusions of the performance of the algorithm depending on the type of function.

Taking into account this reasoning, the test set designed by Eiben and Bäck [EB97b] is very adequate. The test set has several well characterized functions that will allow us to obtain and generalize, as far as possible, the results regarding the kind of function involved. Nevertheless, we have added two functions to the test set with the aim of balancing the number of functions of each kind. These two new functions are the function of Rosenbrock [Ros60] extended to p dimensions and the function of Schwefel [Sch81]; both of them have been widely used in evolutive optimization literature. Table [1](#) shows the expression of each function and a summary of its features: separability, multimodality, and regularity.

A function is multimodal if it has two or more local optima. A function of p variables is separable if it can be rewritten as a sum of p functions of just one variable [Had64]. The separability is closely related to the concept of epistasis or interrelation among the variables of the function. In the field of evolutionary computation, the epistasis measures how much the contribution of a gene to the fitness of the individual depends on the values of other genes.

Non separable functions are more difficult to optimize as the accurate search direction depends on two or more genes. On the other hand, separable functions can be optimized for each variable in turn. The problem is even more difficult if the function is also multimodal. The search process must be able to avoid the regions around local minima in order to approximate, as far as possible, the global optimum. The most complex case appears when the local optima are randomly distributed in the search space.

The dimensionality of the search space is another important factor in the complexity of the problem. A study of the dimensionality problem and its features was carried out by Friedman [Fri94]. In order to establish the same degree of difficulty in all the problems, we have chosen a search space of dimensionality $p = 30$ for all the functions.

Table 1: Definition of each function together with its features

Function	Definition	Multimodal?	Separable?	Regular?
Sphere	$f_{Sph}(\mathbf{x}) = \sum_{i=1}^{p} x_i^2$	no	yes	n/a
	$x_i \in [-5.12, 5.12]$			
	$\mathbf{x}^* = (0, 0, \ldots, 0);\ f_{Sph}(\mathbf{x}^*) = 0$			
Schwefel's	$f_{SchDS}(\mathbf{x}) = \sum_{i=1}^{p} \left(\sum_{j=1}^{i} x_j \right)^2$	no	no	n/a
double sum	$x_i \in [-65.536, 65.536]$			
	$\mathbf{x}^* = (0, 0, \ldots, 0);\ f_{SchDS}(\mathbf{x}^*) = 0$			
Rosenbrock	$f_{Ros}(\mathbf{x}) = \sum_{i=1}^{p-1} [100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2]$	no	no	n/a
	$x_i \in [-2.048, 2.048]$			
	$\mathbf{x}^* = (1, 1, \ldots, 1);\ f_{Ros}(\mathbf{x}^*) = 0$			
Rastrigin	$f_{Ras}(\mathbf{x}) = 10p + \sum_{i=1}^{p} (x_i^2 - 10\cos(2\pi x_i))$	yes	yes	n/a
	$x_i \in [-5.12, 5.12]$			
	$\mathbf{x}^* = (0, 0, \ldots, 0);\ f_{Ras}(\mathbf{x}^*) = 0$			
Schwefel	$f_{Sch}(\mathbf{x}) = 418.9829 \cdot p + \sum_{i=1}^{p} x_i \sin\left(\sqrt{|x_i|}\right)$	yes	yes	n/a
	$x_i \in [-512.03, 511.97]$			
	$\mathbf{x}^* = (-420.9687, \ldots, -420.9687);\ f_{Sch}(\mathbf{x}^*) = 0$			
Ackley	$f_{Ack}(\mathbf{x}) = 20 + e - 20exp\left(-0.2\sqrt{\frac{1}{p}\sum_{i=1}^{p} x_i^2}\right) -$	yes	no	yes
	$exp\left(\frac{1}{p}\sum_{i=1}^{p} \cos(2\pi x_i)\right)$			

	$x_i \in [-30, 30]$			
	$\mathbf{x}^* = (0, 0, \ldots, 0);\ f_{Ack}(\mathbf{x}^*) = 0$			
Griewangk	$f_{Gri}(\mathbf{x}) = 1 + \sum_{i=1}^{p} \frac{x_i^2}{4000} - \prod_{i=1}^{p} cos\left(\frac{x_i}{\sqrt{i}}\right)$	yes	no	yes
	$x_i \in [-600, 600]$			
	$\mathbf{x}^*(0, 0, \ldots, 0);\ f_{Gri}(\mathbf{x}^*) = 0$			
Fletcher	$f_{Flc}(\mathbf{x}) = \sum_{i=1}^{p}(A_i - B_i)^2$	yes	no	no
Powell	$A_i = \sum_{j=1}^{p}(a_{ij}\sin\alpha_j + b_{ij}\cos\alpha_j)$			
	$B_i = \sum_{j=1}^{p}(a_{ij}\sin x_j + b_{ij}\cos x_j)$			
	$x_i, \alpha_i \in [-\pi, \pi];\ a_{ij}, b_{ij} \in [-100, 100]$			
	$\mathbf{x}^* = \alpha;\ f_{Flc}(\mathbf{x}^*) = 0$			
Langerman	$f_{Lan}(\mathbf{x}) = -\sum_{i=1}^{m} c_i \cdot exp\left(-\frac{1}{\pi}\sum_{j=1}^{p}(x_j - a_{ij})^2\right) \cdot$	yes	no	no
	$cos\left(\pi\sum_{j=1}^{p}(x_j - a_{ij})^2\right)$			
	$x_i \in [0, 10];\ m = p$			
	$\mathbf{x}^* = random;\ f_{Lan}(\mathbf{x}^*) = random$			

Sphere function has been used in the development of the theory of evolutionary strategies [Rec73], and in the evaluation of genetic algorithms as part of the test set proposed by De Jong [De 75]. Sphere, or De Jong's function F1, is a simple and strongly convex function. Schwefel's double sum function was proposed by Schwefel [Sch95]. Its main difficulty is that its gradient is not oriented along their axis due to the epistasis among their variables; in this way, the algorithms that use the gradient converge very slowly. Rosenbrock function [Ros60], or De Jong's function F2, is a two dimensional function with a deep valley with the shape of a parabola of the form $x_1^2 = x_2$ that leads to the global minimum. Due to

the non-linearity of the valley, many algorithms converge slowly because they change the direction of the search repeatedly. The extended version of this function was proposed by Spedicato [Spe75]. Other

versions have been proposed [Ore74,Dix74]. It is considered by many authors as a challenge for any optimization algorithm [SV94]. Its difficulty is mainly due to the non-linear interaction among its variables.

Rastrigin function [Ras74] was constructed from Sphere adding a modulator term $\alpha \cdot cos(2\pi x_i)$. Its contour is made up of a large number of local minima whose value increases with the distance to the global minimum. The surface of Schwefel function [Sch81] is composed of a great number of peaks and valleys. The function has a second best minimum far from the global minimum where many search algorithms are trapped. Moreover, the global minimum is near the bounds of the domain.

Ackley, originally proposed by Ackley [Ack87] and generalized by Bäck [BS93], has an exponential term that covers its surface with numerous local minima. The complexity of this function is moderated. An algorithm that only uses the gradient steepest descent will be trapped in a local optima, but any search strategy that analyzes a wider region will be able to cross the valley among the optima and achieve better results. In order to obtain good results for this function, the search strategy must combine the exploratory and exploitative components efficiently. Griewangk function [BFM97] has a product term that introduces interdependence among the variables. The aim is the failure of the techniques that optimize each variable independently. As in Ackley function, the optima of Griewangk function are regularly distributed.

The functions of Fletcher-Powell [FP63] and Langerman [BDL$^{\pm}$96] are highly multimodal, as Ackley and Griewangk, but they are non-symmetrical and their local optima are randomly distributed. In this way, the objective function has no implicit symmetry advantages that might simplify optimization for certain algorithms. Fletcher-Powel function achieves the random distribution of the optima choosing the values of the matrixes **a** and **b**, and of the vector ☐ at random. We have used the values provided by Bäck [Bäc96]. For Langerman function, we have used the values of **a** and **c** referenced by Eiben and Bäck [EB97b].

Domingo 2005-07-11