Finding Multiple Solutions for Multimodal Optimization Problems Using a Multi-Objective Evolutionary Approach

Kalyanmoy Deb
Department of Mechanical Engineering
Indian Institute of Technology, Kanpur
PIN 208016, India
deb@iitk.ac.in

Amit Saha
Kanpur Genetic Algorithms Laboratory
Indian Institute of Technology, Kanpur
PIN 208016, India
asaha@iitk.ac.in

ABSTRACT

In a multimodal optimization task, the main purpose is to find multiple optimal (global and local) solutions associated with a single objective function. Starting with the preselection method suggested in 1970, most of the existing evolutionary algorithms based methodologies employ variants of *niching* in an existing single-objective evolutionary algorithm framework so that similar solutions in a population are de-emphasized in order to focus and maintain multiple distant yet near-optimal solutions. In this paper, we use a completely different and generic strategy in which a single-objective multimodal optimization problem is converted into a suitable bi-objective optimization problem so that all local and global optimal solutions become members of the resulting weak Pareto-optimal set. We solve up to 16-variable test-problems having as many as 48 optima and also demonstrate successful results on *constrained* multimodal test-problems, suggested for the first time. The concept of using multi-objective optimization for solving single-objective multimodal problems seems novel and interesting, and importantly opens further avenues for research.

Categories and Subject Descriptors

G.1.6 [**Optimization**]: Unconstrained Optimization, Constrained Optimization

General Terms

Algorithms, Experimentation

Keywords

Multimodal optimization, multi-objective optimization, NSGA-II, Hooke-Jeeve's exploratory search

1. INTRODUCTION

Single-objective optimization problems are usually solved for finding a single optimal solution, despite the existence of multiple optima in the search space. In the presence of multiple global and local optimal solutions in a problem, an algorithm is usually preferred if it is able to avoid local optimal solutions and locate the true global optimum.

However, in many practical optimization problems having multiple optima, it is wise to find as many optimum points as possible for a number of reasons. First, an optimal solution currently favorable (say, due to availability of some critical resources or satisfaction of some codal principles, or others) may not remain to be so in the future. This would then demand the user to operate at a different solution when such a predicament occur. With the knowledge of another optimal solution for the problem which is favorable to the changed scenario, the user can simply switch to this new optimal solution. Second, the sheer knowledge of multiple optimal solutions in the search space may provide useful insights to the properties of optimal solutions of the problem. A similarity analysis of multiple optimal (high-performing) solutions may bring about useful innovative and hidden principles, similar to that observed in Pareto-optimal solutions in a multi-objective problem solving task [8].

All existing methods in evolutionary computing (EC) literature use an additional *niching* operation in some form or the other for maintaining multiple optimum solutions from one generation to next. In this paper, we suggest a different solution methodology based on a multi-objective formulation of the multimodal optimization problem. In addition to the underlying single-objective function, we introduce an additional second objective, the optimum of which will ensure a unique property followed by every optimal solution in the search space which are not shared by any other solution in the search space. Our approach is in tune with another study [12], which suggested a multi-objectivization technique to convert an otherwise difficult single-objective problem having many local optima into a bi-objective optimization problem. But, our study here is different in the sense that our goal is to find multiple optimal solutions (local and global) by converting the single-objective optimization problem into a bi-objective problem for which all local and global optima are weak Pareto-optimal points.

First we demonstrate a proof-of-principle approach that may be used for problems which are at least twice differentiable. Next, we present a computationally inexpensive approach which can also be used in non-differentiable problems. We also demonstrate the solution of *constrained* multimodal test-problems by making a simple adaptation in our algorithm for unconstrained problems.

In the remainder of the paper, we provide a brief descrip-

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
GECCO'10, July 7–11, 2010, Portland, Oregon, USA.
Copyright 2010 ACM 978-1-4503-0072-8/10/07 ...$10.00.

tion of past multimodal EC studies in Section 2. The concept of multi-objective optimization for solving multimodal problems is described in Section 3. A couple of gradient-based approaches are presented and simulation results on a simple problem are shown to illustrate the principle of multi-objective solution. Thereafter, a number of difficulties in using the gradient approaches in practice are mentioned and a more pragmatic neighborhood based technique is suggested in Section 4. A new Hooke-Jeeves type neighborhood based algorithm is introduced with the proposed multi-objective optimization algorithm in Section 4.3. Results are shown up to 16-variable problems having as many as 48 optima. Thereafter, we discus our results on a constrained multimodal problem in Section 5. Finally, conclusions and a few extensions to this study are highlighted in Section 6.

2. MULTIMODAL EVOLUTIONARY OPTIMIZATION

As the name suggests, a multimodal optimization problem has multiple optimum solutions, of which some can be global optimum solutions having identical objective function value and some can be local optimum solutions having different objective function value. Multimodality in a search and optimization algorithm usually causes difficulty to any optimization algorithm, since there are many attractors, to which the algorithm can become directed to. The task in a multimodal optimization algorithm is to find the multiple optima (global and local) either simultaneously or one after another systematically.

Evolutionary algorithms (EAs) with some changes in the basic framework have been found to be particularly useful in finding multiple optimal solutions simultaneously, simply due to their population approach and their flexibility in modifying the search operators. For making these algorithms suitable for solving multimodal optimization problems, the main challenge has been to maintain an adequate diversity among population members such that multiple optimum solutions can be found and maintained from one generation to another. For this purpose, niching methodologies are employed, in which crowded solutions in the population (usually in the decision variable space) are degraded either by directly reducing the fitness value of neighboring solutions (such as the sharing function approach [2, 5, 10]) or by directly ignoring crowded neighbors (crowding approach [9, 14], clearing approach [15, 21] and others), clustering approach [22] or by other means [13]. A number of recent studies indicates the importance of multimodal evolutionary optimization [6, 7, 21].

The niching methods are also used for other non-gradient methods such as particle swarm optimization [1], differential evolution [11], and evolution strategies [19, 20].

In this study, we take a completely different approach and employ a multi-objective optimization strategy using evolutionary algorithms to find multiple optimal solutions simultaneously. To the best of our knowledge, such a generic approach has not been considered yet in the context of multimodal optimization but our findings indicate that such a multi-objective approach can be an efficient way of solving multimodal unconstrained and constrained problems. Related to our study, there are some studies which used diversity as a second objective. Specifically, [16] describes such a technique in the context of an application problem, but none of these diversity preserving studies attempted to find multiple optima, rather they used the bi-objective optimization to avoid getting stuck to a possible local optimum.

3. MULTIMODAL OPTIMIZATION USING MULTI-OBJECTIVE OPTIMIZATION

In a multimodal optimization problem, we are interested in finding multiple optimal solutions in a single execution of an algorithm. In order to use a multi-objective optimization methodology for this purpose, we need to first identify at least a couple of conflicting (or invariant) objectives for which multiple optimal solutions in a multimodal problem become the trade-off optimal solutions to the corresponding multi-objective optimization problem.

Let us consider the multimodal minimization problem shown in Figure 1 having two minima with different function values:

$$\begin{array}{ll} \text{minimize} & f(x) = 1 - \exp(-x^2)\sin^2(2\pi x), \\ \text{subject to} & 0 \le x \le 1. \end{array} \quad (1)$$

If we arrange the minima according to ascending of their objective function value ($f(x)$), the optimal solutions will line up from the global minimum to the worst local minimum point. In order to have all the minimum points on the trade-off front of a two-objective optimization problem, we need another objective which is either conflicting to $f(x)$ (so they appear on a Pareto-optimal front) or is invariant for all minimum points (so they appear on a weak Pareto-optimal front). We first suggest a couple of gradient-based methods to mainly demonstrate the principle of multi-objectivization idea and then present a couple of neighborhood based approaches which can be used in practice.

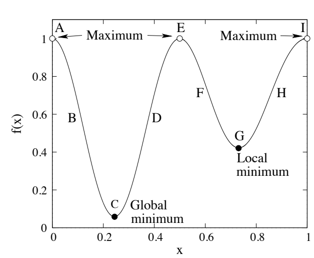

Figure 1: A bi-modal function.

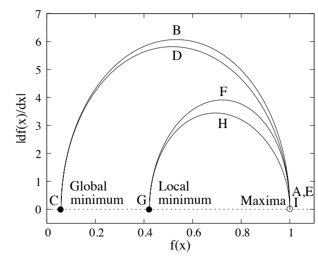

Figure 2: Objective space for minimizing $f(x)$ and $|f'(x)|$.

3.1 Gradient Based Approach

One property which all minimum points will have in common and which is not shared by other points in the search space (except the maximum and saddle points) is that the derivative of the objective function ($f'(x)$) is zero at these points. We shall discuss about a procedure for distinguishing maximum points from the minimum points later[1], but first we explain the multi-objective concept here. Let us consider the following two objectives:

$$\begin{array}{ll} \text{minimize} & f_1(x) = f(x), \\ \text{minimize} & f_2(x) = |f'(x)|, \end{array} \quad (2)$$

[1]We have avoided the cases with saddle points in this gradient approach here, but our latter approaches distinguish minimum points from maximum and saddle points.

in the range $0 \leq x \leq 1$, we observe that the minimum (C and G) and maximum points (A, E and I) of $f(x)$ will correspond to the weak Pareto-optimal points of the above multi-objective problem. Figure 2 shows the corresponding two-dimensional objective space (f_1-f_2). Any point x maps only on the f_1-f_2 curve shown in the figure. Thus, no points other than the points on the shown curve exist in the objective space. It is interesting to observe how f' and f combination makes two different minimum points (C and G) as weak Pareto-optimal solutions of the above multi-objective problem. It is worth reiterating that on the $f_2 = 0$ line, there does not exist any feasible point other than the three points (two corresponding to minimum points and one corresponding to all maximum points) shown in the figure. This observation motivates us to use an EMO procedure in finding all weak Pareto-optimal solutions simultaneously. As a result of the process, we would discover multiple minimum points in a single run of the EMO procedure.

The above multi-objective formulation will also make all maximum points (A, E and I) as weak Pareto-optimal with the minimum points. To avoid this scenario, we can use an alternate second objective, as follows:

$$\begin{array}{ll} \text{minimize} & f_1(x) = f(x), \\ \text{minimize} & f_2(x) = |f'(x)| + (1 - \text{sign}(f''(x))), \end{array} \qquad (3)$$

where sign() returns $+1$ if the operand is positive and -1 if the operand is negative.

For a minimum point, $f''(x) > 0$, so the second term in $f_2(x)$ is zero and $f_2(x) = |f'(x)|$. On the other hand, for a maximum point, $f''(x) < 0$, and $f_2(x) = 2 + |f'(x)|$. For $f''(x) = 0$, $f_2(x) = 1 + |f'(x)|$. This modification will make the maximum and $f''(x) = 0$ points dominated by the minimum points. Figure 3 shows the objective space with this modified second objective function on the same problem considered in Figure 1.

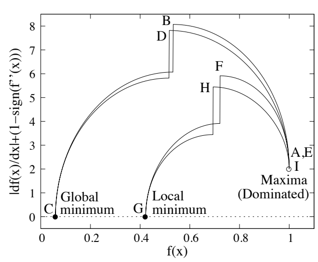

Figure 3: The modified objective space with second-order derivative information makes the maximum points dominated.

3.2 Modified NSGA-II Procedure

The above discussion reveals that the multiple minimum points become different weak Pareto-optimal points of the corresponding multi-objective minimization problem given in equations 2 and 3. However, state-of-the-art EMO algorithms are usually designed to find Pareto-optimal solutions and are not expected to find weak Pareto-optimal solutions. For our purpose here, we need to modify an EMO algorithm to find weak Pareto-optimal points. Here, we discuss the modifications made on a specific EMO algorithm (the NSGA-II procedure [4]) to find weak Pareto-optimal solutions:

1. First, we change the definition of domination between two points a and b. The solution a dominates solution b, if $f_2(a) < f_2(b)$ and $f_1(a) \leq f_1(b)$. Thus, if two solutions have identical f_2 value, they cannot dominate each other. This property will allow two solutions having identical f_2 values to co-survive in the population, thereby allowing us to maintain multiple weak Pareto-optimal solutions, if found in an EMO population.

2. Second, we introduce a *clearing* concept around some selected non-dominated points in the objective space, so as to avoid crowding around minimum points. For this purpose, all points of a non-dominated front (using the above modified domination principle) are first sorted in ascending order of $f(x)$. Thereafter, the point with the smallest $z = f(x)$ solution (having objective values $(f_1, f_2) = (z, z')$) is kept and all population members (including the current non-dominated front members) in the neighborhood around a box ($f_1 \in [z, z+\delta_f]$ and $f_2 \in [z', z'+\delta_{f'}]$ are cleared and assigned a large non-domination rank. The next solution from the sorted $f(x)$ list is then considered and all solutions around its neighborhood are cleared likewise. After all non-dominated front members are considered, the remaining population members are used to find the next non-dominated set of solutions and the above procedure is repeated. This process continues till all population members are either cleared or assigned a non-domination level. All cleared members are accumulated in the final front and solutions with a larger crowding distance value based on $f(x)$ are preferred.

These two modifications ensure that weak non-dominated solutions with an identical $f_2(x)$ values but different $f(x)$ values are emphasized in the population and non-dominated solutions with smaller $f(x)$ values are preferred and solutions around them are de-emphasized.

3.3 Proof-of-Principle Results with the Gradient Approach

In this subsection, we show the results of the modified NSGA-II procedure on a single-variable problem having five minima:

$$f(x) = 1.1 - \exp(-2x)\sin^2(5\pi x), \quad 0 \leq x \leq 1. \qquad (4)$$

Following parameter values are used: population size = 60, SBX probability = 0.9, SBX index = 10, polynomial mutation probability = 0.5, mutation index = 50, $\delta_f = 0.02$, $\delta_{f'} = 0.1$, and maximum number of generations = 100.

Figure 4 shows that the modified NSGA-II is able to find all five minima of this problem. Despite having sparsed points near all five minima, the procedure with its modified domination principle and clearing approach is able to maintain a well-distributed set of points on all five minima.

In addition to finding all the minimum points, the procedure also finds three of the six maximum points (points A, B, and C, shown in the figure). Since the second objective is the absolute value of $f'(\mathbf{x})$, for both minimum and maximum points, it is zero. Thus, all minimum and maximum points are weak dominated points and become the target of above procedure.

In order to avoid finding the maximum points, next, we use the modified second objective function involving the second derivative $f''(x)$, given in equation 3. Identical parameter settings to that in the above simulation are used here. Since the maximum points now get dominated by any minimum point, this time, we are able to cleanly find all the five minima alone (Figure 5). It is also interesting to note that due to the emphasis on finding the weak Pareto-optimal solutions in the suggested multi-objective approach, some solutions near the optima are also found. Since gradient infor-

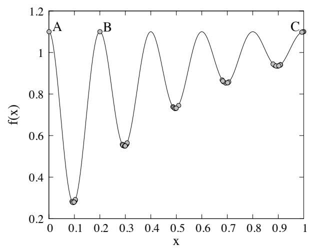

Figure 4: All five minimum points and three maximum points are found by the modified NSGA-II procedure with first derivative information.

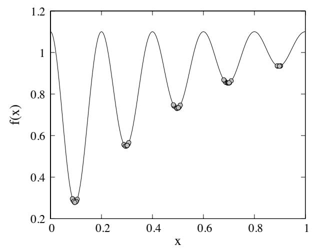

Figure 5: The modified second objective eliminates the maximum points. All five minimum points are found with both first and second derivative information.

mation is used in the second objective, lethal solutions (in non-optimal regions) get dominated by minimum solutions and there is no need for an explicit mating restriction procedure here, which is otherwise recommended in an usual evolutionary multimodal optimization algorithm [5].

3.4 Practical Difficulties with Gradient Based Methods

The gradient based approaches have certain well-known difficulties.

First, such a method has restricted applications, because the method cannot be applied in problems which do not have gradient at some intermediate points or at the minimum points. The gradient information drives the search towards the minimum points and if gradient information is not accurate, such a method is not likely to work well.

Second, the gradient information is specific to the objective function and the range of gradient values near one optimum point may vary from another, depending on flatness of the function values near optimum points. A problem having a differing gradient values near different optima will provide unequal importance to different optima, thereby causing an algorithm difficulty in finding all optimum points in a single run.

Third and a crucial issue is related to the computational complexity of the approach. Although the first-order derivative approach requires $2n$ function evaluations for each population member, this comes with the additional burden of finding the maximum points. To avoid finding unwanted maximum points, the second-derivative approach can be used. But the computation of all second derivatives numerically takes $2n^2 + 1$ function evaluations [17], which is prohibitory for large-sized problems.

The above study is pedagogical and demonstrates that a multimodal optimization problem can be in principle converted to an equivalent multi-objective problem by using first and second-order derivative based optimality conditions, however there are implementation issues which will restrict the use of the gradient-based methods to reasonably higher dimensional problems. Nevertheless, the idea of using a multi-objective optimization technique to find multiple optimum points in a multimodal optimization problems is interesting and next we suggest a more pragmatic approach.

4. MULTIMODAL OPTIMIZATION USING NEIGHBORING SOLUTIONS

Instead of checking the gradients for establishing optimality of a solution ($\mathbf{x}$), we can simply compare a sample of neighboring solutions with the current solution $\mathbf{x}$. The second objective function $f_2(\mathbf{x})$ can be assigned as the count of the number of neighboring solutions which are better than the current solution $\mathbf{x}$ in terms of their objective function ($f(\mathbf{x})$) values.

4.1 NSGA-II with Neighboring Point Count

Only solutions close to a minimum solution will have a zero count on f_2, as locally there does not exist any neighboring solution smaller than the minimum solution. Since f_2 now takes integer values only, the f_2-space is discrete.

To introduce the niching idea on the objective space, we also consider all solutions having an identical f_2 value and clear all solutions within a distance δ_f from the minimum $f(\mathbf{x})$ value. Figure 6 illustrates the modified non-domination procedure. A population of 40 solutions are plotted on a f_1-f_2 space. Three non-dominated fronts are marked in the fig-

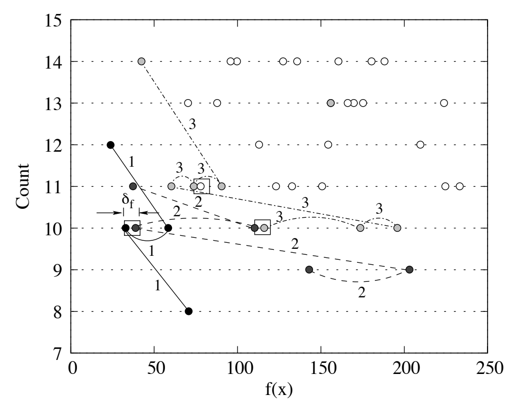

Figure 6: A set of 40 members are ranked according to increasing level of non-domination based on the modified principle and niching operation. Three non-dominated fronts are marked to illustrate the niched domination principle.

ure to make the domination principle along with the niching operation clear. The population has one point with $f_2 = 8$, two points with $f_2 = 9$, seven points with $f_2 = 10$, and so on, as shown in the figure. The point with $f_2 = 8$ dominates (in the usual sense) both $f_2 = 9$ points and the right-most four $f_2 = 10$ points. In an usual scenario, the minimum f_1 point among $f_2 = 10$ points would have dominated other six $f_2 = 10$ points. But due to the modified domination principle, three left-most points with $f_2 = 10$ are non-dominated to each other and are non-dominated with the $f_2 = 8$ point as well. However, the niching consideration makes the second point on $f_2 = 10$ line from left dominated by the first point. This is because the second point is within δ_f distance from the first point. The third point on $f_2 = 10$ line is more than δ_f distance away from the left-most point on $f_2 = 10$ line and hence the third point qualifies to be on the first non-dominated front. Since all points with $f_2 = 11$ get dominated by the left-most point on the $f_2 = 10$ line, none of them qualifies to be on the first non-dominated front. However, the left-most point with $f_2 = 12$ is non-dominated (in the usual sense) with the other first front members. Thus, the best non-dominated front is constituted with one point having $f_2 = 8$, two points having $f_2 = 10$ and one point having $f_2 = 12$. These points are joined a solid line in the figure to show that they belong to the same front.

Similarly, the second non-dominated front members are identified by using the modified domination principle and the niching operation and are marked in the figure with a dashed line. It is interesting to note that these front classification causes a different outcome from that the usual domination principle would have produced. Points lying sparsely in smaller f_2 lines and having smaller f_1 value are emphasized. Such an emphasis will eventually lead the EMO procedure to solutions on $f_2 = 0$ line and the niching procedure will help maintain multiple solutions in the population.

4.2 Proof-of-Principle Results

We consider the following two-variable problem:

$$\begin{array}{ll} \text{min.} & f(x_1, x_2) = (x_1^2 + x_1 + x_2^2 + 2.1x_2) + \sum_{i=1}^{2} 10\,(1 - \cos(2\pi x_i))\,, \\ \text{s.t.} & 0.5 \le x_1 \le (K_1 + 0.5), \quad 0.5 \le x_2 \le (K_2 + 0.5). \end{array} \tag{5}$$

There is a minimum point close to every integer value of each variable within the lower and upper bounds. Since there are K_i integers for each variable within $x_i \in [0.5, K_i + 0.5]$, the total number of minima in this problem are $M = K_1 K_2$.

In our two-objective approach, the first objective is $f(\mathbf{x})$ and the second objective is the number of neighboring points better than $\mathbf{x}$. Figure 7 shows that 97 (out of 100) optima are found by our approach. Figure 8 shows the number of minima found in five problems studied with 10 runs for each problem. The average number of obtained minima are shown with circles and worst and best numbers are shown with bars (which can be hardly seen in the figure). The proposed algorithm is found to be reliably scalable in finding up to 500 optimal points in a single simulation. The modified domination and the niching strategy seems to work together in finding almost all available minima (and as large as 500) for the chosen problems.

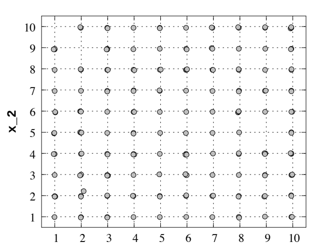

Figure 7: 97 out of 100 minimum points are found by the proposed procedure.

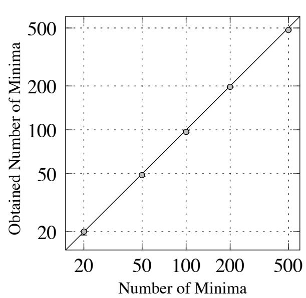

Figure 8: Almost all minimum points are found on two-variable problems by the proposed procedure.

4.3 Neighborhood count with Hooke-Jeeve's Exploratory Search

Instead of generating H solutions at random around a given solution point ($\mathbf{x}$), as done above, we can choose the points judiciously by using the way an exploratory search is performed in the Hooke-Jeeve's optimization algorithm [17]. The procedure starts with the first variable ($i = 1$) dimension and creates two extra points $x_i^c \pm \delta_i$ around the current solution $\mathbf{x}^c = \mathbf{x}$. Thereafter, three solutions $(\mathbf{x}^c - \delta_i \mathbf{e}_i, \mathbf{x}^c, \mathbf{x}^c + \delta_i \mathbf{e}_i)$ (where $\mathbf{e}_i$ is the unit vector along i-th variable axis in the n-dimensional variable space) are compared with their function values and the best is chosen. The current solution $\mathbf{x}^c$ is then moved to the best solution. Similar operations are done for $i = 2$ and continued for the remaining variables. Every time a solution having a better objective value than the original objective value ($f(\mathbf{x})$) is encountered, the second objective value $f_2(\mathbf{x})$ is incremented by one. This procedure requires a total of $H = 2n$ function evaluations to compute $f_2(\mathbf{x})$ for every solution $\mathbf{x}$ in solving an n-dimensional problem. Although this procedure requires a similar amount of additional function evaluations ($2n$) to that in a numerical implementation of the first-derivative approach described earlier, the idea here is more generic and does not have the difficulties associated with the first-derivative approach.

4.4 Results with H-J Approach

The past multimodal evolutionary algorithms studies handled problems having a few variables. In this section, we present results on a number of multi-dimensional multimodal test-problems. First, we define a scalable n-variable test-problem, as follows:

$$\text{MMP}(n): \begin{array}{ll} \text{min.} & f(\mathbf{x}) = \sum_{i=1}^{n} 10(1 + \cos(2\pi k_i x_i)) + 2k_i x_i^2, \\ \text{s.t.} & 0 \le x_i \le 1, \quad i = 1, 2, \ldots, n. \end{array} \tag{6}$$

This function is similar to the Rastrigin's function. Here, the total number of global and local minima are $M = \Pi_{i=1}^{n} k_i$. We use the following k_i values for the three MMP problems, each having 48 minimum points:

$$\begin{array}{lllll} \text{MMP(4):} & k_1 = 2, & k_2 = 2, & k_3 = 3, & k_4 = 4. \\ \text{MMP(8):} & k_2 = 2, & k_4 = 2, & k_6 = 3, & k_8 = 4, \\ & k_i = 1, & \text{for } i = 1, 3, 5, 7. & & \\ \text{MMP(16):} & k_4 = 2, & k_8 = 2, & k_{12} = 3, & k_{16} = 4, \\ & k_i = 1, & \text{for } i = \text{1-3, 5-7, 9-11, 13-15.} & & \end{array}$$

All of these problems have one global minimum and 47 local minimum points. The theoretical minimum variable values are computed by using the first and second-order optimality conditions and solving the resulting root finding problem numerically:

$$x_i - 5\pi \sin(2\pi k_i x_i) = 0.$$

In the range $x_i \in [0, 1]$, the above equation has k_i roots. Different k_i values and corresponding k_i number of x_i variable values are shown in Table 1. In all runs, we use $\delta_i = 0.005n$. For the domination definition, we have used $\delta_f = 0.5$. For all problems, we have used a population size of

Table 1: Optimal x_i values for different k_i values. There are k_i number of solutions for each k_i.

	k_i			
	1	2	3	4
x_i	0.49498	0.24874	0.16611	0.12468
		0.74622	0.49832	0.37405
			0.83053	0.62342
				0.87279

$N = 15\max(n, M)$, where n is the number of variables and M is the number of optima. Crossover and mutation probabilities and their indices are the same as before.

First, we solve the four-variable ($n = 4$) problem. Figure 9 shows the theoretical objective value ($f(\mathbf{x})$) of all 48 minimum points with diamonds. The figure also shows the minimum points found by the proposed H-J based NSGA-II procedure with circles. Only 14 out of 48 minimum points are found. We explain this (apparently poor) performance of the proposed procedure here.

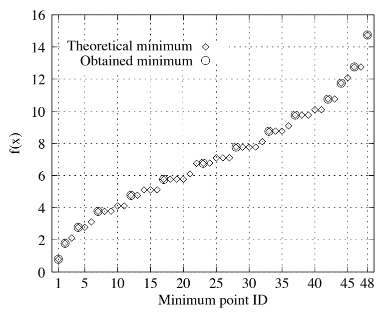

Figure 9: Only 14 of the 48 minimum points are found for the MMP(4) problem by the objective-space niching algorithm.

The figure shows that although there are 48 minima, a number of minima have identical objective values. For example, 14, 15, and 16-th minimum points have identical function value of 5.105. Since the modified domination procedure de-emphasizes all solutions having a difference of $\delta_f = 0.2$ function value from each other, from a cluster of multiple minimum points having identical function value only one of the minimum points is expected to be found; other minimum solutions will be cleared. This scenario can be observed from the figure. Once a point is found, no other minimum point having a similar value is obtained by the algorithm.

4.5 Multiple Global Optimal Solutions

In order to find multiple minimum points having identical or almost equal function values (multiple global minima, for example), we need to perform an additional variable-space niching among closer objective solutions by emphasizing solutions having widely different variable vectors. To implement the idea, we check the normalized Euclidean distance (in the variable space) of any two solutions having function values within δ_f. If the normalized distance is greater than another parameter δ_x, we assign both solutions an identical non-dominated rank, otherwise both solutions are considered arising from the same optimal basin and we assign a large dominated rank to the solution having the worse objective value $f(\mathbf{x})$. In all runs here, we use $\delta_x = 0.2$ here.

Figure 10 summarizes the result obtained by this modified procedure on the 4-variable MMP problem. All 48 minimum points are discovered by the modified procedure in each of the 10 runs started with a different initial random population.

Next, we solve the $n = 16$-variable problem using the modified procedure. Figure 11 shows that the modified procedure is able to find all 48 minimum points in each of the 10 independent runs. There are not many 16-variable multimodal problems that have been attempted in the existing

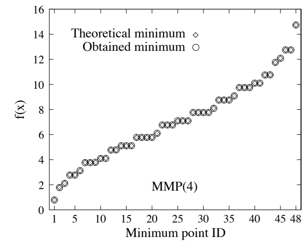

Figure 10: All 48 minimum points are found for the four-variable MMP(4) problem.

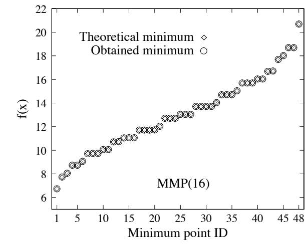

Figure 11: All 48 minimum points are found for the 16-variable MMP(16) problem.

multimodal evolutionary optimization studies. Here, we demonstrate successful working of our procedure on such relatively large-sized problems and also having as many as 48 global minimum points.

Most multimodal EAs require additional parameters to identify niches around each optimum. Our proposed H-J based procedure is not free from additional parameters. However, both δ_f and δ_x parameters directly control the differences between any two optima in objective and decision variable spaces which are desired. Thus, the setting of these two parameters can be motivated from a practical standpoint. However, we are currently pursuing ways of self-adapting these two parameters based on population statistics and the desired number of optima in a problem. Nevertheless, the demonstration of successful working of the proposed approach up to 16-variable problems and having as many as 48 optima amply indicates its efficacy and should motivate further future studies in favor of the proposed multi-objective concept.

5. CONSTRAINED MULTIMODAL OPTIMIZATION

Most studies on evolutionary multimodal optimization concentrated on solving unconstrained multimodal problems and some unconstrained test-problem generators [18] now exist. To our knowledge, a systematic study on suggesting test-problems on *constrained* multimodal optimization problems and algorithms to find multiple optima which lie on constraint boundaries do not exist. However, practical optimization problems most often involve constraints and in such cases multiple optima are likely to lie on one or more constraint boundaries. It is not clear (and has not demonstrated well in the literature yet) whether existing evolutionary multimodal optimization algorithms are adequate to solve such constrained problems.

Here, we suggest a n-variable, J-constraint test-problem. There are J non-convex constraints and the minimum solutions lie on the intersection on the constraint boundaries. We call these problems CMMP(n, G, L), where n is the number of variables, G is the number of global minima and L is the number of local minima in the problem.

PROBLEM CMMP$(n, 2^J, 0)$:

$$\begin{array}{rl} \text{min.} & f(\mathbf{x}) = \sum_{i=1}^{n} x_i^2, \\ \text{s.t.} & g_1(\mathbf{x}) \equiv x_1^2 + 4x_2^2 + 9x_3^2 + \cdots + n^2 x_n^2 \geq n^2, \\ & g_2(\mathbf{x}) \equiv n^2 x_1^2 + x_2^2 + 4x_3^2 + \cdots + (n-1)^2 x_n^2 \geq n^2, \\ & \vdots \end{array} \tag{7}$$

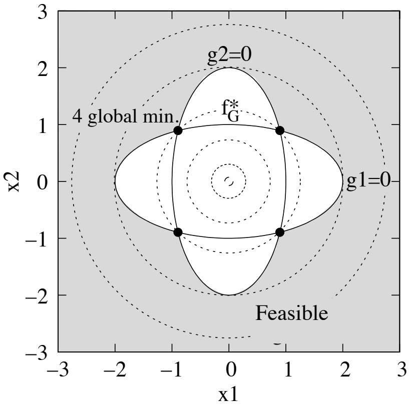

Figure 12: Four global minimum points for the two-variable, two constraint CMMP(2,4,0) problem.

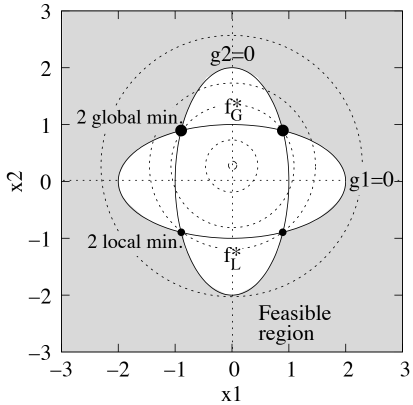

Figure 13: Two global and two local minimum points for the two-variable, two constraint CMMP(2,2,2) problem.

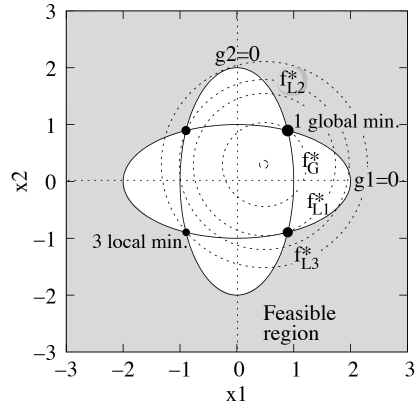

Figure 14: One global and three local minimum points for the two-variable, two constraint CMMP(2,1,3) problem.

$$g_J(\mathbf{x}) \equiv C_{J,1}^2 x_1^2 + C_{J,2}^2 x_2^2 + C_{J,3}^2 x_3^2 + \cdots + C_{J,n}^2 x_n^2 \geq n^2,$$
$$-(n+1) \leq x_i \leq (n+1), \quad \text{for } i = 1, 2, \ldots, n,$$

where

$$C_{j,k} = \begin{cases} (n-j+k+1) \mod n, & \text{if } (n-j+k+1) \mod n \neq 0, \\ n, & \text{otherwise.} \end{cases}$$

Figure 12 shows the constraint boundaries and the feasible search space (shaded region) for $n = 2$ and $J = 2$. Since the square of the distance from the origin is minimized, there are clearly four *global* minima in this problem. These points are intersection points of both ellipses (denoting the constraint boundaries) and are shown in the figure with circles. The optimal points are at $x_1^* = \pm\sqrt{0.8}$ and $x_2^* = \pm\sqrt{0.8}$ with a function value equal to $f^* = 1.6$. Contour lines of the objective function shows that all four minimum points have the same function value.

Interestingly, if we modify the objective function as $f(x_1, x_2) = x_1^2 + (x_2 - 0.2)^2$ (problem CMMP(2,2,2)), there are two global minima: $(\sqrt{0.8}, \sqrt{0.8})$, $(-\sqrt{0.8}, \sqrt{0.8})$ with a function value $f_G^* = 1.282$ and there are two local minima: $(\sqrt{0.8}, -\sqrt{0.8})$, $(-\sqrt{0.8}, -\sqrt{0.8})$ with the function value $f^* = 1.998$. Figure 13 shows that the contour of the objective function starting with a value zero at $(0, 0.2)$ and then increasing with a point's distance from $(0, 0.2)$. The global and local minima are also marked in the figure.

If the objective function is modified as $f(x_1, x_2) = (x_1 - 0.3)^2 + (x_2 - 0.2)^2$ (the problem CMMP(2,1,3)), there are three different local minima and one global minimum, as shown in Figure 14. The global minimum is at $(\sqrt{0.8}, \sqrt{0.8})$ with a function value $f^* = 0.836$. The three local minima are $(\sqrt{0.8}, -\sqrt{0.8})$ with a function value $f^* = 1.551$, $(-\sqrt{0.8}, \sqrt{0.8})$ with a function value $f^* = 1.909$ and $(-\sqrt{0.8}, -\sqrt{0.8})$ with a function value $f^* = 2.624$.

5.1 Constrained Handling Procedure for Multimodal Optimization

To take care of feasible and infeasible solutions present in a population, we use the constraint handling strategy of the NSGA-II procedure [3, 4]. When two solutions are compared for domination, one of the following three scenarios can happen: (i) when one solution is feasible and other is infeasible, we choose the feasible solution, (ii) when two infeasible solutions are being compared, we simply choose the solution having smaller overall normalized constraint violation. and (iii) when both solutions are feasible, we follow the niching comparison method used before for unconstrained problems. Instead of using a procedure which is completely different from the unconstrained method, we use a simple modification to our proposed approach and importantly without the need of any additional constrained handling parameter. To get an accurate evaluation of f_2 for solutions close to a constraint boundary, we ignore any infeasible solution in the count of better neighboring solutions. The H-J exploratory search is used for evaluating the second objective function.

5.2 Constraint Handling Results

In this subsection, we present our results on the above problems and attempt to solve them using the proposed procedure. Identical NSGA-II parameter values as those used in the unconstrained cases are used here. For all constrained problems, we use $\delta_f = 0.2$ and $\delta_x = 0.08$. We perform 10 runs and show results from a typical run.

Table 2 presents the four minimum points found by the proposed procedure on the three problems. By comparing with Figures 12, 13 and 14, it can be observed that the obtained points are identical to the four minimum points.

Figure 15 shows the obtained solutions on the 10-variable, four-constraint CMMP(10,16,0) problem. In this problem, the intermediate six variables (x_4 to x_9) take a value close to zero and 16 solutions come from each of two values of other four variables. The objective function values among 16 solutions vary only a maximum of 0.43% in the range [1.614, 1.621].

6. CONCLUSIONS

In this paper, we have suggested a multi-objective formulation of a multimodal optimization problem so that multiple global and local optimal solutions become the only candidate weak Pareto-optimal set. While the objective function of the multimodal optimization problem is one of the objectives, a number of suggestions of the second objective function have been made here. Starting with the gradient-based approaches (demonstrating the foundation of the multi-objective approach), more pragmatic neighborhood count based approaches have been systematically developed for this pur-

Table 2: Multimodal points found by constraint handling procedure for different problems.

Problem		Soln. 1	Soln. 2	Soln. 3	Soln. 4
CMMP(2,4,0)	$\mathbf{x}$	(0.894, 0.895)	(-0.895, 0.895)	(0.894, -0.896)	(-0.894, -0.897)
	f	1.601	1.601	1.602	1.603
CMMP(2,2,2)	$\mathbf{x}$	(0.894, 0.895)	(0.895, -0.894)	(0.894, -0.895)	(-0.894, -0.896)
	f	1.583	1.584	1.619	1.620
CMMP(2,1,3)	$\mathbf{x}$	(0.896, 0.894)	(-0.897, 0.894)	(0.894, -0.895)	(-0.894, -0.895)
	f	1.548	1.587	1.619	1.655

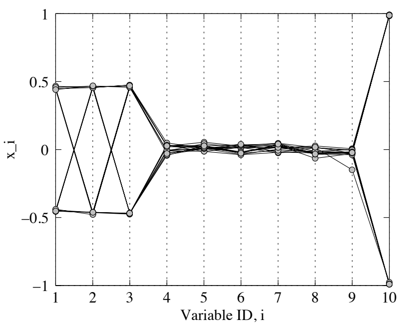

Figure 15: 16 global minimum points for the 10-variable, four-constraint CMMP(10,16,0) problem.

pose. Modifications to an existing EMO procedure are made in order to find weak non-dominated solutions. Using the Hooke-Jeeves based neighborhood count method for the second objective, the proposed EMO procedure has been able to solve 16-variable, 48-optima problems having a combination of global and local optima. Another hallmark of this study is the suggestion and solution of a multimodal constrained test-problem generator which is scalable in terms of the number of variables, constraints, and optima.

The multimodal solution methodology suggested here is different from the usual niching approaches and borrows ideas from the evolutionary multi-objective optimization. Further efforts must now be spent to evaluate and extend the ideas to more complex problems.

Acknowledgments

Authors wish to thank Prof. Juergen Branke for sharing some initial thoughts with the first author during his Bessel research award visit to University of Karlsruhe in 2003.

References

[1] J. Barrera and C. A. C. Coello. *A Review of Particle Swarm Optimization Methods Used for Multimodal Optimization*, pages 9–37. Berlin: Springer, 2009.

[2] D. Beasley, D. R. Bull, and R. R. Martin. A sequential niche technique for multimodal function optimization. *Evolutionary Computation*, 1(2):101–125, 1993.

[3] K. Deb. *Multi-objective optimization using evolutionary algorithms*. Wiley, Chichester, UK, 2001.

[4] K. Deb, S. Agrawal, A. Pratap, and T. Meyarivan. A fast and elitist multi-objective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, 6(2):182–197, 2002.

[5] K. Deb and D. E. Goldberg. An investigation of niche and species formation in genetic function optimization. In *Proceedings of the Third International Conference on Genetic Algorithms*, pages 42–50, 1989.

[6] K. Deb, J. Horn, and D. E. Goldberg. Multi-modal deceptive functions. *Complex Systems*, 7(2):131–153, 1993.

[7] K. Deb and A. Kumar. Real-coded genetic algorithms with simulated binary crossover: Studies on multi-modal and multi-objective problems. *Complex Systems*, 9(6):431–454, 1995.

[8] K. Deb and A. Srinivasan. Innovization: Innovating design principles through optimization. In *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO-2006)*, pages 1629–1636, New York: ACM, 2006.

[9] K. A. DeJong. *An Analysis of the Behavior of a Class of Genetic Adaptive Systems*. PhD thesis, Ann Arbor, MI: University of Michigan, 1975. Dissertation Abstracts International 36(10), 5140B (University Microfilms No. 76-9381).

[10] D. E. Goldberg and J. Richardson. Genetic algorithms with sharing for multimodal function optimization. In *Proceedings of the First International Conference on Genetic Algorithms and Their Applications*, pages 41–49, 1987.

[11] Z. Hendershot. A differential evolution algorithm for automatically discovering multiple global optima in multidimensional, discontinuous spaces. In *Proceedings of 15th Midwest Artificial Intelligence and Cognitive Sciences Conference*, pages 92–97, 2004.

[12] J. D. Knowles, R. A. Watson, and D. Corne. Reducing local optima in single-objective problems by multi-objectivization. In *Proceedings of the Evolutionary Multi-Criterion Optimization Conference (EMO 2001)*, pages 269–283, 2001.

[13] J.-P. Li, M. E. Balazs, G. T. Parks, and P. J. Clarkson. A species conserving genetic algorithm for multimodal function optimization. *Evolutionary Computation*, 10(3):207–234, 2002.

[14] O. Mengsheol and D. E. Goldberg. Probabilistic crowding: deterministic crowding with probabilistic replacement. In *Proceedings of genetic and Evolutionary computation conference (GECCO-1999)*, pages 409–416, 1999.

[15] A. Pétrowski. A clearing procedure as a niching method for genetic algorithms. In *IEEE 3rd International Conference on Evolutionary Computation (ICEC'96)*, pages 798–803, 1996.

[16] M. Preuss, G. Rudolph, and F. Tumakaka. Solving multimodal problems via multiobjective techniques with Application to phase equilibrium detection. In *Proceedings of the Congress on Evolutionary Computation*. Piscataway, NJ: IEEE Press, 2007.

[17] G. V. Reklaitis, A. Ravindran, and K. M. Ragsdell. *Engineering Optimization Methods and Applications*. New York : Wiley, 1983.

[18] J. Rönkkönen, X. Li, V. Kyrki, and J. Lampinen. A generator for multimodal test functions with multiple global optima. In *Proceedings of the Seventh International Conference on Simulated Evolution and Learning (SEAL-08)*, pages 239–248, 2008.

[19] O. Shir and T. Bäck. Niching in evolution strategies. In *Proceedings of the Conference on Genetic and Evolutionary Computation (GECCO-2005)*, pages 915–916, 2005.

[20] O. Shir and T. Bäck. Niche radius adaptation in the CMA-ES niching algorithm. In *Parallel Problem Solving from Nature (PPSN-9)*, pages 142–151, 2006.

[21] G. Singh and K. Deb. Comparison of multi-modal optimization algorithms based on evolutionary algorithms. In *Proceedings of the 8th annual conference on Genetic and evolutionary computation, July*, pages 08–12, 2006.

[22] F. Streichert, G. Stein, H. Ulmer, and A. Zell. A clustering based niching ea for multimodal search spaces. In *Proceedings of the International Conference Evolution Artificielle*, pages 293–304, 2003. LNCS 2936.