A basis for future successes in multiobjective combinatorial optimization

Pedro Castro Borges

and

Michael Pilegaard Hansen

Institute of Mathematical Modelling

Technical University of Denmark

Building 321

2800 Lyngby

Denmark

Email: {pmb, mph}@imm.dtu.dk

URL: http://www.imm.dtu.dk/~{pmb, mph}

March 27, 1998

IMM Technical Report IMM-REP-1998-8

A basis for future successes in multiobjective combinatorial optimization

Pedro Castro Borges and Michael Pilegaard Hansen

Institute of Mathematical Modelling, Technical University of Denmark, 2800 Lyngby, Denmark, email: {pmb, mph}@imm.dtu.dk

Abstract: This paper addresses global convexity in multiobjective combinatorial optimization: a phenomenon that has previously been studied and shown to exist in a number of single objective combinatorial problems. Global convexity can imply that local optima are concentrated in a very small region of the solution space when distance is measured according to the topology of a neighborhood function. Global convexity is therefore not an absolute characteristic of an optimization problem, but refers of the topology that has been induced in the solution space. Local search based heuristics can exploit global convexity by choosing neighborhood functions that tend to concentrate the search in these regions and it has been claimed that global convexity plays an important role in the recent successes of many metaheuristics.

This report generalizes and extends those investigations for the case where multiple objectives exist by analyzing several points where global convexity might play a role. This analysis builds on generating local optima for well known scalarizing functions, covering the entire set of weights. We study the heuristic concentration of local optima, the distance between local optima for different weights over the non-dominated frontier and the stability of local optima for small weight variations. If global convexity can be established over local areas of the non-dominated frontier, it must also be exploitable in new metaheuristics for MOCO optimization.

We consider the 2-OPT neighborhood function on a three objective traveling salesman problem and use both the weighted sums program and the augmented Tchebycheff program. The results indicate that there indeed exists global convexity and foresee the development of new local search heuristics for approximating the non-dominated frontier in MOCO.

Keywords: Multiple objective combinatorial optimization; Global convexity; Local search

Introduction

The recent successes of metaheuristics in single objective combinatorial optimization have ignited a rapid growth in their adaptation to multiple objective problems. Such problems have higher degree of intractability than their single objective counterparts and metaheuristics light the hope of handling them. Besides important aspects of implementation and flexibility, the reasons for the success of metaheuristics lie in the quality of the results, which might be due to the existence of *global convexity* in many combinatorial optimization applications. Expressions like *heuristic concentration* and *massif central* have also been used to describe the phenomenon.

Global convexity is not convexity in the strict sense (Hu et. al., 1989) but may be used to explain the empirical observation that most of the best feasible solutions to a combinatorial problem often are gathered in an extremely small fraction of the entire solution space. Metaheuristics may exploit this by concentrating the search in that area.

Any assessment of global convexity for a particular problem only makes sense once a suitable topology has been established the solution space. This may not be straightforward for a given combinatorial problem. A popular topology construction technique is based on neighborhood mapping, which consists in defining a set of neighboring solutions to each solution. Once a neighborhood mapping (also denoted a neighborhood function, or neighborhood structure) is established we can derive a metric in the solution space based on this mapping.

Global convexity has been studied and shown for e.g. the single objective traveling salesman problem (Boese, 1996). In this paper we use an approach based on the one of Boese et. al. (1994) but will extend it to consider the traveling salesman problem with multiple objectives. The aim is to investigate global convexity over regions of the non-dominated frontier by examining whether (exact or potentially) non-dominated solutions which are close in the objective space or in a supporting weight space also are close according to a solution space metric. When and if global convexity exists, it may be exploited in metaheuristics for multiobjective combinatorial optimization.

Section 1 summarizes some basic concepts of multiobjective combinatorial optimization and of local search, including a description of the greedy randomized descent heuristic. Section 2 presents the multiobjective traveling salesman problem. The following four sections present the computational results. Section 3 examines a set of well dispersed optimal supported non-dominated solutions in order to see whether global convexity exists for that sub-set of the non-dominated set. In Section 4 we essentially repeat the experiments of Boese et. al. (1994) on local optima. We use a fixed weight and consider both the weighted sums program and the augmented Tchebycheff program. A generalization of this investigation is presented in Section 5, where the entire weight-space is taken into account. In Section 6 we take a closer look at a smaller local area in the weight-space to assess the stability of the local optima with small weight deviations. Before a brief conclusion, Section 7 offers directions for further research and in particular, how global convexity can be exploited in (new) adaptations of well-known metaheuristics to a multiobjective environment.

1. Basic concepts

1.1 The multiple objective combinatorial optimization (MOCO) problem

The general MOCO problem can be formulated as:

$$\text{"min"} \qquad f(x) = \begin{cases} f_1(x) = z_1 \\ f_2(x) = z_2 \\ \quad \cdots \\ f_J(x) = z_J \end{cases}$$

$$\text{subject to} \quad x \in S \tag{1}$$

where **x** is the decision vector, or *solution*, and *S* is a finite set of feasible solutions, the *solution space*. *J* is the number of objectives and the objective function f(**x**) maps S into $\Re^J$. The *image* of a solution in the decision space is a *point*, $\mathbf{z} = [z_1, z_2, ..., z_J]$. A point, **z**, is *attainable* if there exists a solution $\mathbf{x} \in S$ so that $\mathbf{z} = f(\mathbf{x})$. The set of all attainable point is denoted **Z**.

The *ideal point*, $\mathbf{z}^*$, is defined as $\mathbf{z}^* = [\min f_1(\mathbf{x}), \min f_2(\mathbf{x}), ..., \min f_J(\mathbf{x})]$. Normally, the ideal point is not attainable and there will therefore not exist a single optimal solution to the MOCO problem. Instead one must choose among *non-dominated* attainable points, where an attainable point, **z**, is non-dominated if there does not exist another attainable point **z'** which is better on at least one objective without being worse on others, i.e. where $z'_j \leq z_j$ on all objectives *j*.

A multicriteria decision making (MCDM) process may involve obtaining one, some or all of the non-dominated points. For such tasks some *achievement scalarizing functions* have been established of which we shall focus on two: weighted sums and augmented Tchebycheff scalarizing functions.

The weighted sums scalarizing function can be written as follows:

$$s_{ws}(\mathbf{z}, \lambda) = \sum_{j=1}^{J} \lambda_j (z_j - z^*_j)$$

where $\lambda = [\lambda_1, \lambda_2, ..., \lambda_J]$ is a *weight vector* with non-negative elements and $\mathbf{z}^*$ is the ideal point.

The (weighted) augmented Tchebycheff scalarizing function can be written as:

$$s_{aTch}(\mathbf{z}, \lambda) = \max_j \left\{ \lambda_j (z_j - z^*_j) \right\} + \varepsilon\, s_{ws}(\mathbf{z}, \lambda)$$

where ε is a very small (sufficiently small) positive scalar.

Achievement scalarizing functions are used to project the ideal point onto the non-dominated frontier through achievement *scalarizing programs*. The weighted sums program is as follows:

$$\min s_{ws}(\mathbf{z}, \lambda)$$

$$\text{subject to } \mathbf{z} \in \mathbf{Z} \tag{2}$$

The weighted sums program is often used in practice. But whereas all optima to (2) are non-dominated points, there might exist non-dominated points for which no weight vector exists so that the point minimizes (2). Optimal points to the weighted sums program are called *supported* non-dominated points.

The (weighted) augmented Tchebycheff program is defined as:

$$\begin{aligned} &\min s_{aTch}(\mathbf{z}, \lambda) \\ &\text{subject to } \mathbf{z} \in \mathbf{Z} \end{aligned} \qquad (3)$$

The augmented Tchebycheff program is useful because all non-dominated points can be found as optimum to (3) when an adequate weight vector is used just like any optimum of (3) will be a non-dominated point. For more on scalarizing functions, please refer to Wierzbicki (1986).

Let the weight set Λ be defined as $\Lambda = \{\lambda \in \Re^J \mid \forall_j\ \lambda_j \geq 0 \wedge \lambda_1 + \lambda_2 + \ldots + \lambda_J = 1\}$. We will limit ourselves to consider weight vectors in this set, so $\lambda \in \Lambda$ and measure the distance between two weight vectors, λ^1 and λ^2, by their Manhattan distance:

$$d(\lambda^1, \lambda^2) = \|\lambda^1 - \lambda^2\|_1 = \sum_{j=1}^{J} |\lambda^1{}_j - \lambda^2{}_j|$$

Manhattan distance is also used to measure distance between points in the objective space:

$$d(\mathbf{z}^1, \mathbf{z}^2) = \|\mathbf{z}^1 - \mathbf{z}^2\|_1 = \sum_{j=1}^{J} |z^1{}_j - z^2{}_j|$$

A set $\Lambda' \subset \Lambda$ is *maximally dispersed* (with respect to the Manhattan norm) if $\min d(\lambda^i, \lambda^k) \geq \delta$ for all $\lambda^i, \lambda^k \in \Lambda'$, $i \neq k$, and there does not exist any (larger) set $\Lambda'' \subset \Lambda$ such that $|\Lambda''| > |\Lambda'|$ and $\min d(\lambda^n, \lambda^m) \geq \delta$ for all $\lambda^n, \lambda^m \in \Lambda''$, $n \neq m$

A maximally dispersed set can be constructed from any positive integer *n*, by including all the different weight vectors whose components are multiples of 1/n and which sum up to one. This set will have $\min d(\lambda^i, \lambda^k) = 2/n$, for $i \neq k$ and the number of unique weight vectors is equal to $C(J+n-1, n) = (n+J-1)! \,/\, (n!\,(J-1)!)$.

1.2 Local Search

A *neighborhood structure* is defined by an assignment of a set of neighboring solutions to each solution. The replacement of a solution with a solution from its set of neighboring solutions is a *neighborhood move*. *Local search* performs neighborhood moves in the pursuit of improvement and is a mechanism used by many metaheuristic search strategies, such as steepest descent, simulated annealing, tabu search, and many others.

The set of neighboring solutions to solution $\mathbf{x}$ is denoted by $N(\mathbf{x})$. Solution $\mathbf{x}$ is a *local optimum*, if none of the neighboring solutions are better, i.e. if $f(\mathbf{y}) \geq f(\mathbf{x})$ for all $\mathbf{y} \in N(\mathbf{x})$. Solution $\mathbf{x}$ is a *global optimum* if $f(\mathbf{z}) \geq f(\mathbf{x})$ for any feasible solution $\mathbf{z}$. A simple local search strategy, which locates a local optimum and which will be used in the experiments later on for the scalarizing programs, is the greedy randomized descent as outlined in Figure 1.

For a local search based heuristic to perform well for practical global optimization problems, the neighborhood structure must normally fulfill several criteria, many of which are problem specific. One criterion, however, that seems to be of general importance is that the neighborhood structure does not have a tendency to create deceptive cost surfaces. A deceptive cost surface is created if the global optimum is located far from the local optima. To evaluate this characteristic a distance measure in the solutions space is therefore required.

1. Generate a random starting solution, **y**
2. Repeat
3.   Set **x** equal to **y**
4.   Set *M* equal to N(**x**)
5.   Repeat
6.   Set **y** to a random solution from *M* and remove **y** from *M*
7.   Until f(**y**) < f(**x**) or *M* is empty
8. Until f(y) ≥ f(x)
9. Return local optimal solution, **x**

Figure 1. Greedy randomized descent heuristic

The distance from solutions $\mathbf{x}^1$ to solution $\mathbf{x}^2$ can be measured as the minimum number of neighborhood moves needed to move from one to the other, i.e. as follows: If $\mathbf{x}^1 = \mathbf{x}^2$ then $d(\mathbf{x}^1, \mathbf{x}^2) = 0$, if $\mathbf{x}^2 \in N(\mathbf{x}^1)$ then $d(\mathbf{x}^1, \mathbf{x}^2) = 1$, else $d(\mathbf{x}^1, \mathbf{x}^2)$ is defined as one plus the length of the shortest list of solutions $(\mathbf{y}^1, \mathbf{y}^2, \ldots, \mathbf{y}^k)$, such that $\mathbf{y}^1 \in N(\mathbf{x}^1)$, $\mathbf{y}^i \in N(\mathbf{y}^{i-1})$, $i=2..k$ and $\mathbf{x}^2 \in N(\mathbf{y}^k)$. If the neighborhood structure is symmetric, i.e. if $\mathbf{y} \in N(\mathbf{x}) \Leftrightarrow \mathbf{x} \in N(\mathbf{y})$, then $d(\mathbf{x}^1, \mathbf{x}^2) = d(\mathbf{x}^2, \mathbf{x}^1)$.

2. Multiobjective TSP

We will consider the multiobjective traveling salesman problem (moTSP):

$$
\text{"min"} \qquad f(\mathbf{x}) = \begin{Bmatrix} f_1(\mathbf{x}) = c_1^{\mathbf{x}(n),\mathbf{x}(1)} + \sum_{i=1}^{n-1} c_1^{\mathbf{x}(i),\mathbf{x}(i+1)} \\ f_2(\mathbf{x}) = c_2^{\mathbf{x}(n),\mathbf{x}(1)} + \sum_{i=1}^{n-1} c_2^{\mathbf{x}(i),\mathbf{x}(i+1)} \\ \ldots \\ f_J(\mathbf{x}) = c_J^{\mathbf{x}(n),\mathbf{x}(1)} + \sum_{i=1}^{n-1} c_J^{\mathbf{x}(i),\mathbf{x}(i+1)} \end{Bmatrix} \tag{4}
$$

$$
\text{subject to} \quad \mathbf{x} \in \wp\{1,2,\ldots,n\}
$$

where n is the number of cities, c_j is the cost matrix with respect to objective j, and the decision variable $\mathbf{x}$ holds a permutation of the n cities.

An often used neighborhood structure for the single objective TSP is the 2-edge exchange from the 2-OPT heuristic (Lin, 1965) and consists of removing two edges from the tour and replacing them with the only two other edges that yield a feasible tour. The set of all such neighboring solutions to solution $\mathbf{x}$ is N($\mathbf{x}$).

In the experiments, we optimize on the scalarizing programs (2) and (3) for the moTSP. Notice, however, that while each weighted sums program (2) will be a single objective TSP, program (3) will not, since the objective function is not linear.

For the experiments we chose a 100 cities moTSP with 3 objective, $J = 3$. As data we use some of the Krolak instances with 100 cities from TSPLIB (Reinelt, 1995), namely kroA100, kroB100 and kroC100. Each cost matrix then corresponds to one objective in (6) so that kroA100 corresponds to the cost matrix of objective number 1, kroB100 to objective number 2, and kroC100 to objective number 3. One advantage of using these cost matrices is that the tours can be presumed to be within the same scale of range for the different objectives, making range scaling unnecessary. Also, the data are easily available and the optimal value to each problem is known, which gives us the exact ideal point $\mathbf{z}^*$.

As a surrogate measure for the distance between two solutions we will measure the similarity, or bonds, of two solutions by the number edges common for the two solutions. The bond distance as used by Kirkpatrick and Toulouse (1985) is defined by *n* minus the number of edges common to both solutions. Boese (1996) has developed a dynamic programming procedure to calculate the number of solutions at a given bond distance to an arbitrarily chosen solution. Using their software we computed these numbers for a 100 cities TSP, as shown in Figure 2.

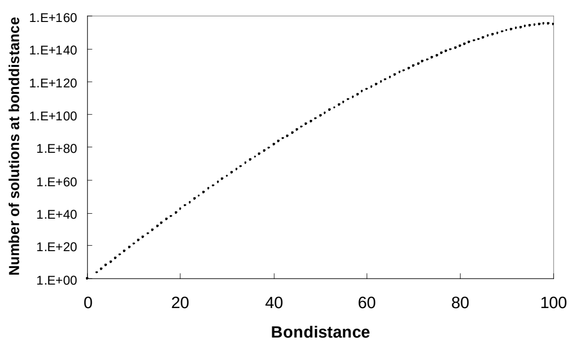

Figure 2. Number of solutions at different bond distances.

The average bond distance is around 98 and the total number of solutions is around $4.66{\cdot}10^{155}$. Within spheres of bond distances 20, 30, 40 and 50 are located fractions of the total number of solutions of $4.61{\cdot}10^{-114}$, $8.82{\cdot}10^{-94}$, $9.89\ {\cdot}10^{-75}$, $4.31\ {\cdot}10^{-57}$, respectively. These figures give an idea of how large a proportion of the solution space actually are outside a “ball” with a given center and given “radius”, as measured in bond distance.

3. Optimal solutions to the weighted sums program

First we take a look at some optimal solutions to the weighted sums program (2). Using n = 100 we generated the C(102, 100) = 5.151 maximally dispersed weight vectors each defining a 100 cities single objective TSP. With the software package tsp_solve (Hurwitz, 1994) the problem instances were solved to optimality, giving 2.150 unique points and (as it turned out) unique solutions. Unfortunately, we are not aware of algorithms for solving the augmented Tchebycheff problem to optimality within reasonable time limits. The optimal solutions to the weighted sums problems are primarily to be used in the experiments of Section 5, but let us here describe some observed phenomena.

Each optimal solution is characterized by a generating weight vector λ, a solution $\mathbf{x}$ in the decision space, and the point $\mathbf{z} = f(\mathbf{x})$ in objective space. Figure 3 shows a contour plot of the frequencies of $d(\lambda^i, \lambda^k)$ and $d(\mathbf{x}^i, \mathbf{x}^k)$ considering all pairs of located optimal solutions $(\mathbf{x}^i, \mathbf{x}^k)$ from all 5.151 weight vectors. It can be seen that optimal solutions belonging to close weight vectors also are located close in the solutions space. Also, the bond distances are surprisingly small, even for weights as different as 0.25 the global optima are generally within bond distances of 50. Only when the weights are completely different, corresponding to different single objective TSP instances, does the bond distance reach 100.

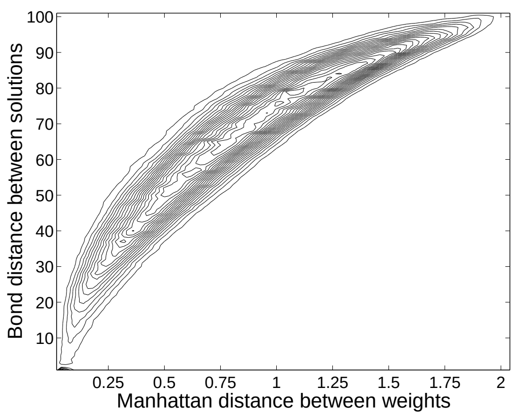

Figure 3. Difference in *weights* vs. difference in solutions for pairs of optimal solutions to the weighted sums program.

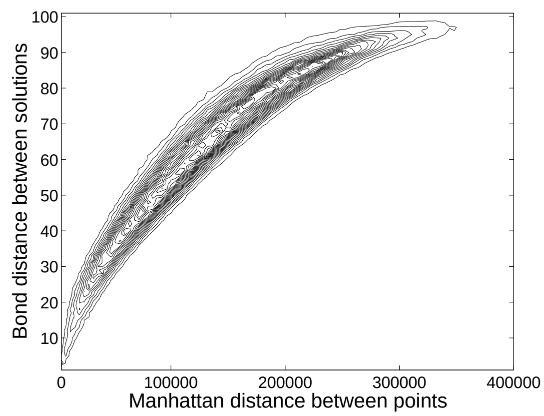

Figure 4. Difference in *points* vs. difference in solutions for pairs of optimal solutions to the weighted sums program.

In Figure 4 the frequencies of $d(\mathbf{z}^i, \mathbf{z}^k)$ and $d(\mathbf{x}^i, \mathbf{x}^k)$ are shown in a contour plot, considering all pairs of the 5.151 located optimal solutions $(\mathbf{x}^1, \mathbf{x}^k)$. We see how optimal solutions that are close on the non-dominated frontier of the (non-weighted) objective space also are close in the decision space. This is often important for the performance of multiobjective heuristics.

4. Heuristic concentration of local optima

Finding the optimal solutions in the previous section was very time-consuming and could not be done for the augmented Tchebycheff program. Often one has to use heuristics to approach problems which are larger or more difficult or for which sophisticated exact algorithms have not been developed, as it has been for the single objective TSP.

To simulate such a situation and to be able to consider also the augmented Tchebycheff program, we use the greedy randomized descent heuristic with the 2-OPT neighborhood structure to locate local optima. Our aim is now to investigate the distribution of those local optima in the decision space for a particular weight vector. The experiments in this Section are based on work by Boese, Kahng and Muddu (1994).

We use one fixed weight vector of (1/3, 1/3, 1/3) and generate 1000 local optima to both the weighted sums problem and the augmented Tchebycheff problem. The optimal solution to the weighted sums problem is also calculated.

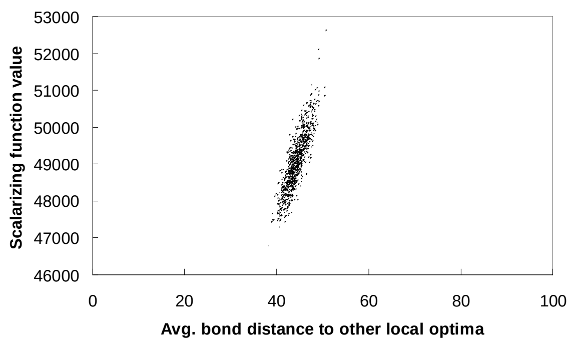

Figure 5. Scalarizing function values of 1000 local optima to the weighted sums program as a function of their average bond distance to the other 999 local optima.

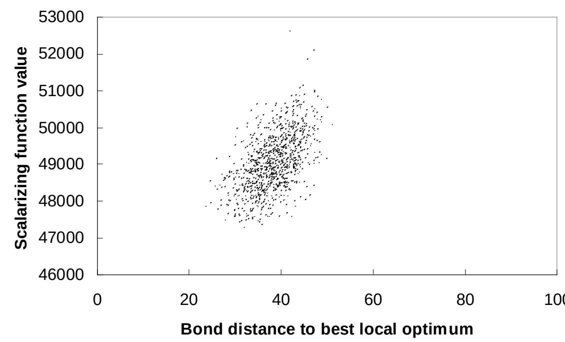

Figure 6. Scalarizing function values of 1000 local optima to the weighted sums program as a function of their average bond distance to the best local optima.

For the weighted sums problem, Figure 5 shows for each local optimum the scalarizing function value as a function of the average bond distance to the other 999 local optima. Figure 6 shows for each local optimum the bond distance to the best of the local optima.[1]

Figures 7 and 8 show analogous graphs for the augmented Tchebycheff problem.

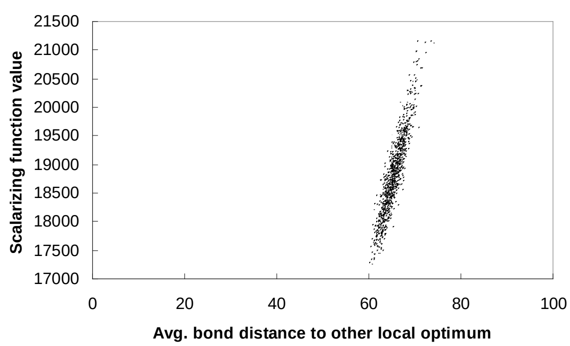

Figure 7. Scalarizing function values of 1000 local optima to the augmented Tchebycheff program as a function of their average bond distance to the other 999 local optima.

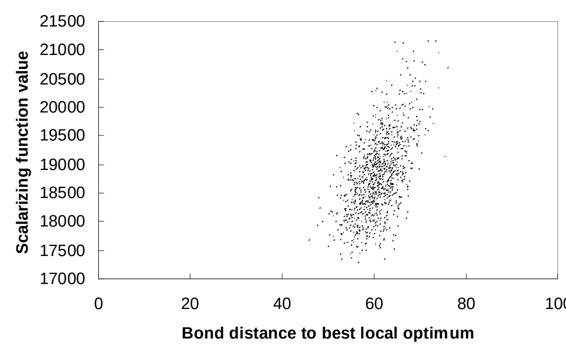

Figure 8. Scalarizing function values of 1000 local optima to the augmented Tchebycheff program as a function of their average bond distance to the best local optima.

First it should be noticed that the bond distances again are rather small considering the vastness of the solution space, and that this is most evident for the weighted sums problem. Secondly we see that local optima, which are closer to other local optima, also are better as opposed to those which are not so centrally located. In particular does the *best* local optimum have an average bond distance to other local optima of 38.3 (60.9) which is the lowest (6th lowest)

---

[1] Note that for Figures 6, 8 and 9 we have to the bond distance added a random number between -0.5 and 0.5 so the points do not describe lines at the original integer values but allow the reader to assess the intensity of the concentrations.

value among the 1000 local optima to the weighted sums (augmented Tchebycheff) program. It looks like there indeed exists global convexity, or heuristic concentration, also for the Tchebycheff program, even is the bond distances here are significantly larger.

In global optimization it is also important, that the local optima are not located too far from the global optimum. For the weighted sums problem, Figure 9 shows for each local optimum the scalarizing function value as a function of the bond distance to the *optimal* solution. We see that the graph is similar to that of Figure 6 and the global optimum likewise seems to be located centrally among the local optima. Also, the *best* local optimum is close to the global optimum as compared to the other local optima.

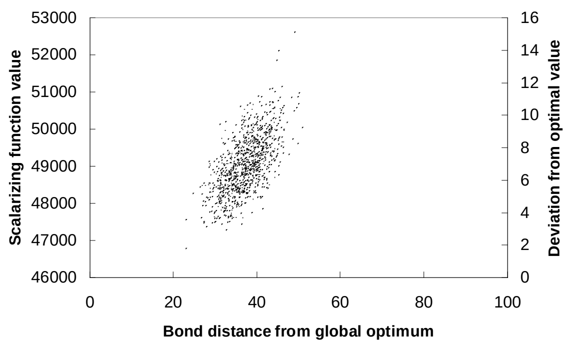

Figure 9. Scalarizing function values of 1000 local optima to the weighted sums program (and their deviation from the optimal value) as a function of their bond distance to the global optimum.

5. Local optima over the entire non-dominated frontier

The previous Section analyzed the dispersion of local optima for only one weight vector. In this Section we extend the analysis to cover the whole set of weights. One thousand local optima were generated for each of the 5.151 maximally dispersed weight vectors. This was done both for the weighted sums program and for the augmented Tchebycheff program.

First we will take a look at whether the results of Section 4 are general for the whole non-dominated frontier. The convexity assumption would say that the local optima are gathered in a relatively small region of the solution space. The centrality assumption says further that the *best* of these local optima are located central with respect to the other local optima. If both assumptions are valid, we should also expect that if the local optima are gathered in smaller regions, then these should also be closer to the best local optimum. That this holds is seen in Figures 10 and 11.

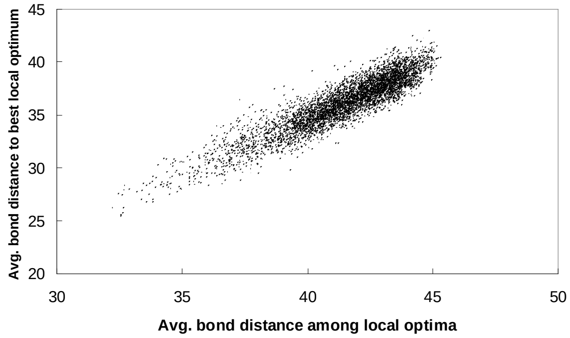

Figure 10. For each of the 5.151 experiments for the weighted sums program is plotted the relationship between the avg. bond distance among the local optima and avg. bond distance to the best local optimum.

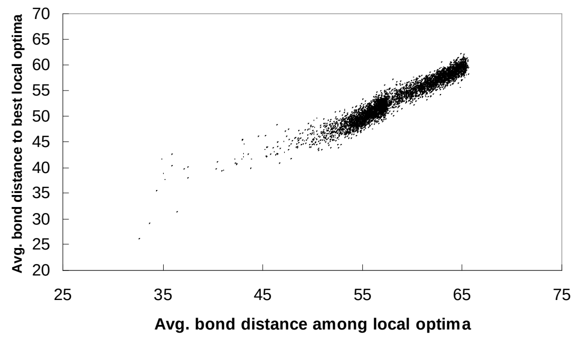

Figure 11. For each of the 5.151 experiments for the augmented Tchebycheff program is plotted the relationship between the avg. bond distance among the local optima and avg. bond distance to the best local optimum.

From the global optimization point of view it would be nice if we could expect that if the local optima are gathered in small regions, then these are also close to the *global* optimum. Figure 12 shows that this indeed is the case for the local optima of the weighted sums programs.

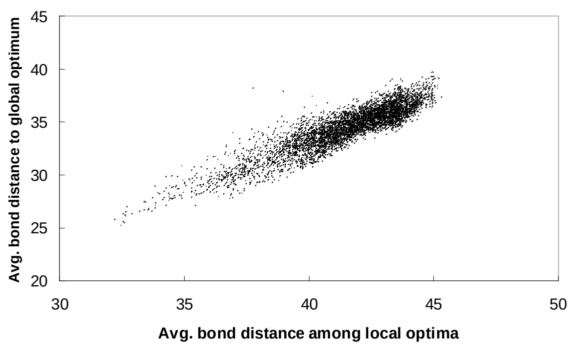

Figure 12. For each of the 5.151 experiments for the weighted sums program is plotted the relationship between the avg. bond distance among the local optima and avg. bond distance to the global optimum.

Another interesting aspect is analyzing how often the best local optima for the weighted sums problem were closest local optima to the global optimum. The proportion of best local optima which were the solutions closest to the respective global optima was 18% (of the 5.151 cases). In over 90% of the cases the best local optima were in the 10% local optima closest to the respective global optimum. This indicates that in general the best of the local optima is closer to the global optimum.

In Section 3 we looked at the relationships between distances in weights, solutions and points for the global optima of the weighted sums program. We will now consider the same kind of contour plots, but for the best local optima both of the weighted sums program and of the Tchebycheff program. These are represented in Figures 13 through 16. For the weighted sums program we see a similar pattern to the one observed for the optimal solutions, only the bond distances start at a higher value. The plots for the best local optima for the augmented Tchebycheff program show the expected higher bond distances. The higher spread of the Tchebycheff heuristic might be originated only by the poorer performance of the heuristic itself. One can not exclude that if we had other heuristics for the Tchebycheff program then we would have results more similar to the ones found for the weighted sums.

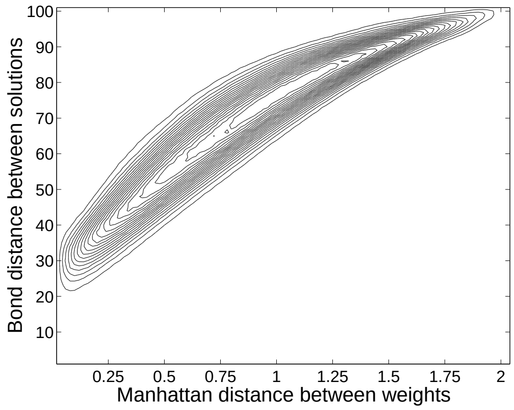

Figure 13. Difference in *weights* vs. difference in solutions for pairs of best local solutions to the weighted sums program.

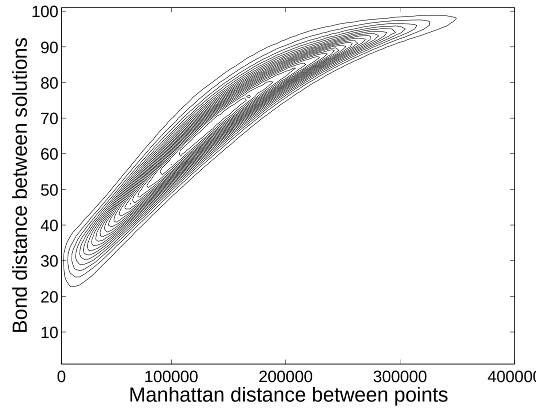

Figure 14. Difference in *points* vs. difference in solutions for pairs of best local solutions to the weighted sums program.

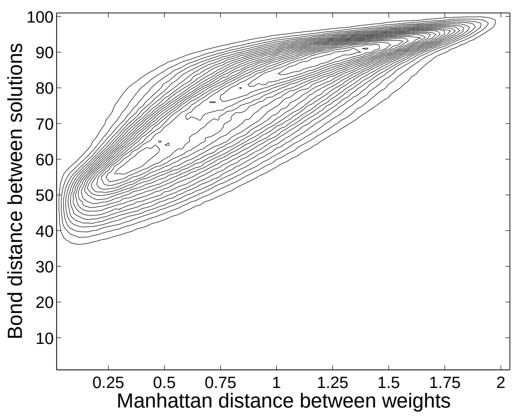

Figure 15. Difference in *weights* vs. difference in solutions for pairs of best local solutions to the augmented Tchebycheff program.

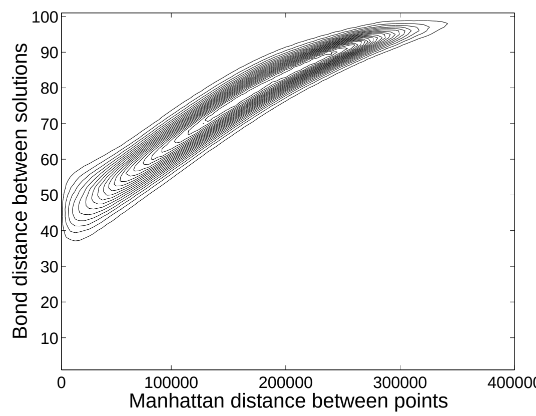

Figure 16. Difference in *points* vs. difference in solutions for pairs of best local solutions to the augmented Tchebycheff program.

6. Analysis of a local area of the non-dominated frontier

In this Section we intend to analyze the impacts of small changes in the weights vector. Such analysis is relevant since some kind of weight modifications is common in many MOCO procedures. Our approach consists in zooming in on a smaller area around a particular weight vector by generating well spread weight vectors at a given distance of that *central vector*. The local optima found with the spread weight vectors can afterwards be compared to the best local optimum (out of 1000 runs) found with the central vector.

Again we chose to zoom in on the area around the central weight (1/3, 1/3, 1/3). To be able to examine what happens for different weight differences the nearby weights were generated for 100 different distances equally spread from 0.003 to 0.300. The weights at a fixed distance, or “radius”, from the central point measured with the Manhattan distance describe a hexagonal line in the weigh space when projected on the 2D plane. A total of 1200 equidistant weights vectors were generated for each fixed distance, which corresponds to having 200 equidistantly spread weights on each line segment of that hexagon.

For each weigh vector one local optimum was found for both the weighted sums program and for the augmented Tchebycheff program.

We consider the scalarizing function values of the local optima averaged over the 1200 weights vector with the same distance to the central weight. Figures 17 and 18 show the results and it seems clear that the scalarizing function values are rather constant for the weighted sums program when the weight vector only is changed marginally. This is not so for the augmented Tchebycheff program and it may indicate, that the weighted sums program is more robust with respect to uncertain or imprecise weights and therefore is easier to handle in practice.

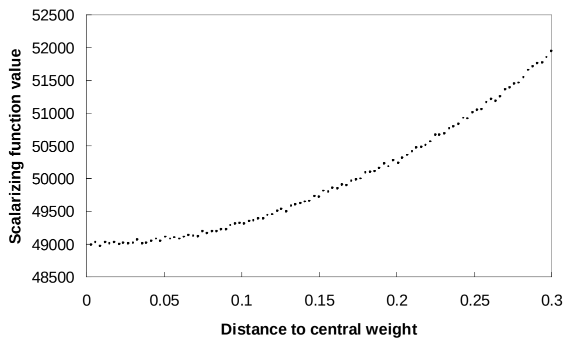

Figure 17. Average scalarizing function values for local optima at different distances to the central point in the weighted sums program.

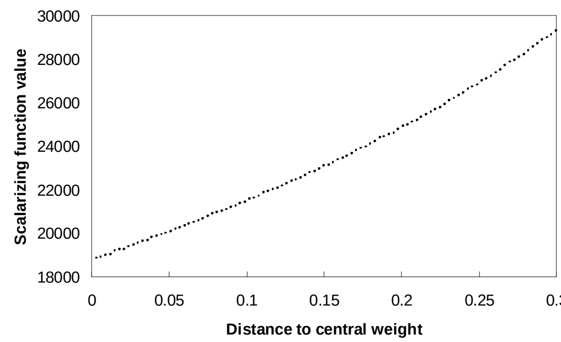

Figure 18. Average scalarizing function values for local optima at different distances to the central point in the augmented Tchebycheff program.

In Figures 19 and 20 the reverse is seen when we observe the distance in the objective space between the best local optima at the central weight (as found in Section 4) and the local optima with weights at small distances from the central weight. But we also notice that the weighted sums programs start at a smaller distance and that it constantly keeps a lower level.

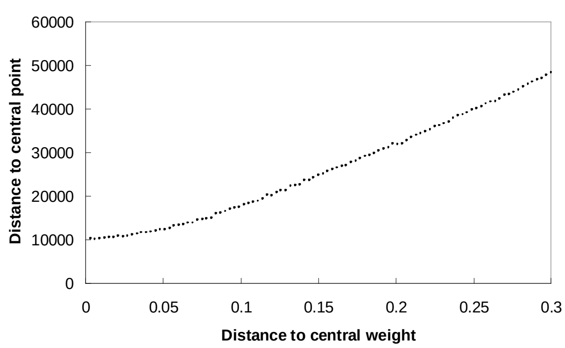

Figure 19. Average distance in objective space from local optima at given distances from the central weight to the best local optimum for the central weight for the weighted sums program.

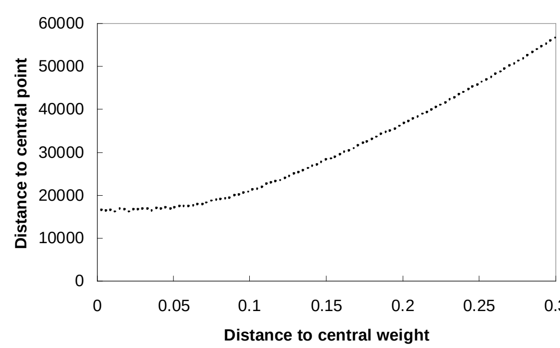

Figure 20. Average distance in objective space from local optima at given distances from the central weight to the best local optimum for the central weight for the augmented Tchebycheff program.

It is important to remember, however, that the weights are used in different programs so that the influence of small changes of the weights does not represent a direct comparison.

The experiments of this Section gave the same general outcome when it was repeated with three other central weights, namely (0.4, 0.3, 0.3), (0.3, 0.4, 0.3) and (0.3, 0.3, 0.4), but with fewer local optima for each distance to the central weight.

7. Implications and directions for future research

The conclusions of this paper apply to the problem and problem instance used herein, even if we believe from our experience with MOCO problems they are more general. Naturally it will be beneficial to perform the presented type of investigation on instances with correlated objectives since they are very common in practice. Other MOCO problems should also be considered.

Special attention should be given to MOCO problems of mixed structures, that is, MOCO problems where the objective functions belong to different problem classes. The MOCO problem of this paper is a TSP in all objectives, and the 2-OPT is recognized as a suitable neighborhood structure for this problem. If some objectives were not of the TSP structure, the 2-OPT may not be as well chosen and the global convexity not as obvious or non-existent.

Our results are based on observed behavior rather than theoretical analysis. More theoretical work could be beneficial in the field. Especially on the difference and similarities in the behaviors of the weighted sums problem and the Tchebycheff scalarizing function.

Most interesting would be the development of MOCO metaheuristics that can exploit global convexity whenever it exists. Instead of making direct translations of single objective metaheuristics in the sense of repeatedly optimizing scalarizing functions of different weights, one could try to use previously obtained information gained from the solution structure of obtained good solutions close (in the objective space or in the weight space) to the current one. This is perhaps most easily implemented with heuristics which consider not one, but a multitude of current solutions, as genetic algorithms (Goldberg, 1989) and scatter search (Glover, 1995).

Genetic algorithms combine known solutions into new solutions that are located “between” the known solutions. The term “between” is here based on the topology induced in the solution space as defined by the crossover operator. Whether the traditional building block viewpoint or the solution space topology viewpoint is more relevant depends on the purpose but it can be read from the work presented here, that the latter certainly has a usage in MOCO heuristics. Hybrid genetic algorithms as well as scatter search based heuristics adapted for MOCO problems may have a bright future and experiments with these is an obvious path for future research. Indeed, genetic algorithms have already been widely used for multiobjective problems (survey by Fonseca and Fleming, 1995).

But also the traditional single solution metaheuristics like simulated annealing and tabu search (Glover 1989, 1990) may be adapted for MOCO. Multiple solution adaptations for MOCO problems have been presented by Czyzak and Jaszkiewicz (1998) and Hansen (1997a) for simulated annealing and tabu search, respectively. These methods could also be adapted to exploit global convexity.

Greedy randomized adaptive search procedures or GRASP (Feo and Resende, 1995), heuristic concentration (Rosing and ReVelle, 1997) and jump search (Tsumakitani and Evans, 1998) are already prepared for exploiting global convexity, only they have not been adapted for MOCO problems. We see such an adaptation as very valuable for the field of MOCO metaheuristics. In particular, they may be used in single solutions MOCO approaches such as NISE (Cohon, 1978), MOSA (Ulungu et. al., 1995), MOTS (Gandibleux et. al., 1996) and CHESS (Borges, 1997). This last approach explores the existing information by optimizing a function of the previously found points and focusing only on solution and objective spaces instead of weight space, not being bound to a particular direction at a time.

Conclusion

The main result of this paper is to show existence of global convexity for a multiobjective combinatorial optimization problem when using a common neighborhood structure. To be more precise, the experiments conducted in this paper indicate that heuristic concentration is present in a persistent manner over the entire area both of the exact (supported) and heuristically approximated non-dominated frontier. We have shown how the non-dominated (or potentially non-dominated) solutions that result from the weighted sums and Tchebycheff programs vary with the possible weights. This is observed in the solutions space, where distance is measured by a bond distance, and in the objective space.

The association of those distances is far greater than we had anticipated before starting the experiments. We expected that the bond distances would go over 90 after only minor changes in the weights. On the contrary we observed bond distances under 50 for the weighted sums program and under 70 for the augmented program for fairly large changes in the weights. These distances correspond to very small fractions of the solutions space.

The results seem to suggest that the augmented Tchebycheff program is harder to optimize than the weighted sums program, on the moTSP using 2-OPT local optima. Hansen (1997b) found the same result in a tabu search implementation based on 2-OPT on a moTSP problem contained in the problems of this paper. It must be stressed, however, that the results are based on one moTSP using the 2-OPT neighborhood structure. Other problems and other neighborhood structures could lead to other conclusions[2]. In particular, the 2-OPT neighborhood search was used due to its previous successes in single objective TSP as for instance the weighted sums program. The augmented Tchebycheff program is different in nature and the 2-OPT might not be as effective for this program.

As a spin-off from the exhaustive computational experiments, we have obtained a well-dispersed set of supported non-dominated solutions and two set of potentially non-dominated solutions. These sets may serve as reference sets for experiments on approximation of the non-dominated frontier for this moTSP instance.

It is our opinion, that many metaheuristics can be adapted to MOCO problems and, in cases where global convexity has a strong presence, this can and should be exploited in these adaptations. To say it in a another way: metaheuristics for MOCO should be developed so that global convexity can be exploited.

References

Andersen, K. A., Jörnsten, K., Lind, M. (1996), “On bicriterion minimal spanning trees: An approximation”, *Computers and Operations Research*, Vol. 23, No. 12, pp. 1171-1182.

Boese, K. D., Kahng, A. B. and Muddu, S. (1994), “A new adaptive multi-start technique for combinatorial global optimizations”, *Operations Research Letters*, Vol. 16, pp. 101-113.

Boese, K. D. (1996), *Models for iterative global optimization*, Ph.D. dissertation at the Computer Science Department of UCLA.

Borges, P. C. (1997), “CHESS: Changing horizon efficient set search”, submitted to *Journal of Heuristics*.

---

[2] An illustrative example from Andersen et. al. (1996) shows how all supported solutions to the bi-objective spanning tree problem are connected with a minimal bond distance (of one edge) whereas this is not true for neighboring unsupported solutions.

Cohon, J. L. (1978), *Multiobjective programming and planning*, Academic Press.

Czyzak, P. and Jaszkiewicz, A. (1998), "Pareto simulated annealing - a metaheuristic technique for multiple objective combinatorial optimization", to be published in *Journal of Multicriteria Decision Making*

Feo, T. A., Resende, M. G. C. (1995), "Greedy randomized adaptive search procedures", *Journal of Global Optimization*, Vol. 6, pp. 109-133.

Fonseca, C. M. and Fleming, P. J. (1995), "An overview of evolutionary algorithms in multiobjective optimization", *Evolutionary Computation*, Vol. 3, pp. 1-16.

Gandibleux, X., Mezdaoui, N., Fréville, A. (1996), "A tabu search procedure to solve multiobjective combinatorial optimization problems", to appear in *Proceedings volume of MOPGP '96*, R. Caballero and R. Steuer (eds.), Springer-Verlag.

Gardiner, L. R., Steuer, R. E. (1994), "Unified interactive multiple objective programming", *European Journal of Operational Research*, Vol. 74, pp. 391-406.

Glover, F. (1989), "Tabu Search - Part I", *ORSA Journal on Computing*, Vol. 1, No. 3, pp. 190-206.

Glover, F. (1990), "Tabu Search - Part II", *ORSA Journal on Computing*, Vol. 2, No. 1, pp. 4-32.

Glover, F. (1995), "Scatter search and star-paths: beyond the genetic metaphor", *OR Spectrum*, Vol. 17, pp. 125-137.

Goldberg, D. E. (1989), *Genetic algorithms in search, optimization and machine learning*, Addison-Wesley

Hansen, M. P. (1997a), "Tabu search for multiobjective optimization: MOTS", paper presented at the *13th International Conference on Multiple Criteria Decision Making*, University of Cape Town, January 6-10, 1997.

Hansen, M. P. (1997b), "Experiments on the optimization of the Tchebycheff program", submitted to *Journal of Heuristics*.

Hurwitz, C. (1994), The *GNU tsp_solve software package*, version 1.3.6, Free Software Foundation.

Kirkpatrick, S. and Toulouse, G. (1985), "Configuration space analysis of traveling salesman problems", *Journal de Physique*, Vol. 46, pp. 1277-1292.

Hu, T. C., Klee, V., Larman, D. (1989), "Optimization of globally convex functions", *SIAM Journal of Control and Optimization*, Vol. 27, pp. 1026-1047.

Lin, S. (1965), "Computer solutions to the traveling salesman problem", *Bell System Technical Journal*, Vol. 44, pp. 2245-2269.

Reinelt, G. (1995), *TSPLIB*, available at URL http://www.iwr.uni-heidelberg.de/iwr/comopt/soft/TSPLIB95/TSPLIB.html

Rosing, K. E. and ReVelle, C.S. (1997), "Heuristic concentration: two stage solution construction", *European Journal of Operational Research*, Vol. 97, pp. 75-86.

Serafini, P. (1987), "Some considerations about computational complexity for multiobjective combinatorial problems", *Lecture Notes in Economics and Mathematical Systems*, Vol. 294, pp. 222-232.

Serafini, P. (1992), "Simulated annealing for multi objective optimization problems", Proceedings of the Xth International conference on MCDM, Tapei 19-24.07.92, pp. 87-96.

Steuer, R. E. (1986), *Multiple criteria optimization: Theory, computation, and application*, John Wiley & Sons.

Tsumakitani, S. and Evans, J. R. (1998), "An empirical study of a new metaheuristic for the traveling salesman problem", *European Journal of Operational Research*, Vol. 104, pp. 113-128.

Ulungu, E. L. and Teghem, J. (1994), "Multi-objective combinatorial optimization problems: a survey", *Journal of multiple-criteria decision analysis*, Vol. 3, pp. 83-104.

Ulungu, E. L., Teghem and J., Fortemps, Ph. (1995), "Heuristics for multi-objective combinatorial optimization by simulated annealing", *Multiple Criteria Decision Making: Theory and applications. Proceedings of the 6th International Conference on MCDM*, Beijing, China, pp. 228-238.

Wierzbicki, A. P. (1986), "On the completeness and constructiveness of parametric characterizations to vector optimization problems", *OR Spektrum*, Vol. 8, pp. 73-87.

This work was partially supported by grant CIENCIA / BD / 2510