23

Adaptive Formation of Pareto Front in Evolutionary Multi-objective Optimization

Özer Ciftcioglu and Michael S. Bittermann
Delft University of Technology
The Netherlands

1. Introduction

Optimization is an important concept in science and engineering. Traditionally, methods are developed for unconstrained and constrained single objective optimization tasks. However, with the increasing complexity of optimization problems in the modern technological real world problems, multi-objective optimization algorithms are needed and being developed. With the advent of evolutionary algorithms in the last decades, multi-objective evolutionary algorithms (MOEAs) are extensively being investigated for solving associated optimization problems, e.g. (Deb et al., 2000, Zitzler & Thiele, 1999, Yao et al., 1999, Eiben et al., 1999). An updated survey of Ga-based MOEAs is given by (Coello, 1999). Evolutionary algorithms are particularly suitable for this, since they evolve simultaneously a population of potential solutions. These solutions are investigated in non-dominated solution space so that the optimized solutions in a multi-objective functions space form a front which is known as Pareto surface or front. Obtaining this simultaneous solution front in a single run is an appealing property that it is the incentive for a fast growing interest on MOEAs in the last decade. Although Pareto front is an important concept, its formation is not straightforward since the strict search of non-dominated regions in the multi-objective solution space prematurely excludes some of the potential solutions that results in an aggregated solutions in this very space. This means Pareto surface is not fully developed and the diversity of the solutions on the Pareto front is not fully exercised. Conventionally, non-dominated solutions with many objectives are usually low in number making the selection pressure toward the Pareto front also low, with aggregated solutions in the Pareto dominance-based MOEA algorithms (Sato, 2007). The purpose of this research is to investigate this issue and provide effective solutions with fast convergence together with diversity of solutions is maintained on the Pareto front. This goal has already attracted attention in the literature (Laumanns et al., 2002). This work addresses this issue with a novel concept of adaptive formation of Pareto front. This is demonstrated with an application from the domain of architectural design. The method is based on relaxed dominance domains, which basically refer to a degree of relaxation of the dominance in the terminology of MOEAs. In this book-chapter contribution, the relaxed dominance concept is explicitly described and applied. The organisation of this chapter is as follows. Section two describes the relaxed dominance concept. Section three describes the adaptive formation of Pareto front in a design application. This is followed by the conclusions in section four.

Source: Evolutionary Computation, Book edited by: Wellington Pinheiro dos Santos,
ISBN 978-953-307-008-7, pp. 572, October 2009, I-Tech, Vienna, Austria

2. Design computation subject to multiobjective optimization

Multi-objective optimization deals with optimization where several objectives are involved. These objectives are conflicting or in competition among themselves. For a single objective case there are traditionally many algorithms in continuous search space, where gradient-based algorithms are most suitable in many instances. In discrete search spaces, in the last decade evolutionary algorithms are ubiquitously used for optimization, where genetic algorithms (GA) are predominantly applied. However, in many real engineering or design problems, more than two objectives need to be optimized simultaneously. To deal with multi-objectivity it is not difficult to realize that evolutionary algorithms are effective in defining the search direction. Basically, in a multi-objective case the search direction is not one but may be many, so that during the search a single preferred direction cannot be identified. In this case a population of candidate solutions can easily hint about the desired directions of the search and let the candidate solutions during the search process be more probable for the ultimate goal. Next to the principles of GA optimization, in MO algorithms, in many cases the use of Pareto ranking is a fundamental selection method. Its effectiveness is clearly demonstrated for a moderate number of objectives, which are subject to optimization simultaneously (Deb, 2001). Pareto ranking refers to a solution surface in a multidimensional solution space formed by multiple criteria representing the objectives. On this surface, the solutions are diverse but they are assumed to be equivalently valid. The eventual selection of one of the solutions among those many is based on some so-called higher order preferences, which require more insight into the problem at hand. This is necessary in order to make more refined decisions before selecting any solution represented along the Pareto surface.

In solving multi-objective optimization, the effectiveness of evolutionary algorithms has been well established. For this purpose there are quite a few algorithms which are running quite well especially with low dimensionality of the multidimensional space (Coello et al., 2003). However, with the increase of the number of objective functions, i.e. with high dimensionality, the effectiveness of the evolutionary algorithms is hampered. One measure of effectiveness is the expansion of Pareto front where the solution diversity is a desired property. For this purpose, the search space is exhaustively examined with some methods, e.g. *niched Pareto ranking*, e.g. (Horn et al., 1994). However these algorithms are rather involved so that the search needs extensive computer time for a satisfactory solution in terms of a Pareto front. Because of this extensive time requirement, distributed computing of Pareto-optimal solutions is proposed (Deb et al., 2003), where multiple processors are needed. They basically share the computational task with cooperation among each other, making the task scalable (Hughes, 2005, Jaszkiewicz, 2004).

The issue of solution diversity and effective solution for multi-objective optimization problem described above is especially the due to elimination of many acceptable solutions during the evolutionary computation process, in case orthogonal standard Pareto dominance is used. This is a kind of Greedy algorithm which considers the solutions at the search area delimited by orthogonal axes of the multidimensional space. To increase the pressure pushing the Pareto surface towards to the maximally attainable solution point is the main problem and relaxation of the orthogonality with a systematic approach is needed. By such a method next to non-dominated solutions also some dominated solutions are considered at each generation. Such dominated solutions can be potentially favourable solutions in the present generation, so that they can give birth to non-dominated solution in

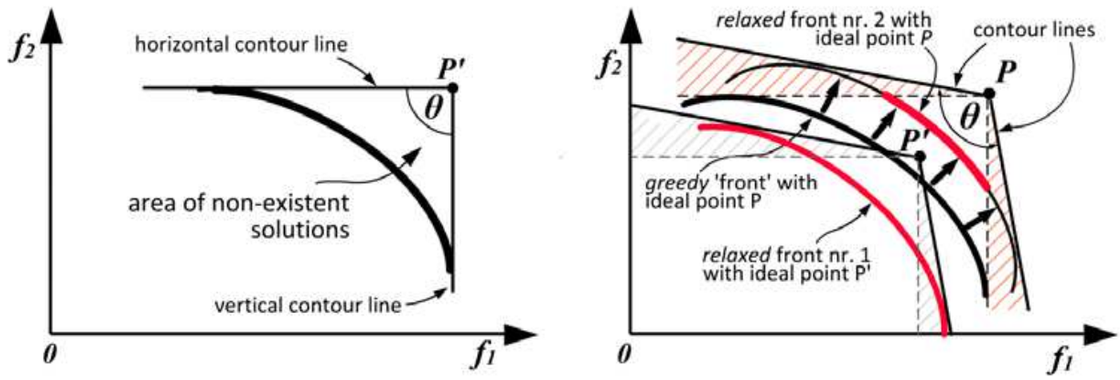
the following generation. Although, some relaxation of the dominance is addressed in literature (Branke et al., 2000, Deb et al., 2006), in a multidimensional space, to identify the size of relaxation corresponding to a volume is not explicitly determined. In such a volume next to non-dominated solutions, dominated but potentially favourable solutions, as described above, lie. To determine this volume optimally as to the circumstantial conditions of the search process is a major and a challenging task. The solution for this task is essentially due to the mathematical treatment of the problem where the volume in question is identified adaptively during the search that it yields a measured pressure to the Pareto front toward to the desired direction, at each generation. In the adaptive process reported in this work, the volume is determined by genetic search for each member of the population. The process is adaptive, because the Pareto front is converged progressively in the course of consecutive generations, where the rate of convergence is determined with volume size, which is subject to appropriate change at each generation. In non-adaptive case, the Pareto front is also converged progressively; however the rate of convergence, in contrast to the adaptive case, is monotonically exhausted. The adaptation is explained shortly afterwards below via contour lines in the objective-functions space. Here the volume with dominated solutions is termed as *relaxed dominance region* and this novel concept is *preliminarily* introduced before (Ciftcioglu & Bittermann, 2008) for a non-adaptive case.

Some important features of the latest generation MOEAs address the selection of the potential solutions during the optimization process, and diversity-preserving strategies in objective space. Next to the principles of GA optimization, in MO algorithms, in many cases the use of Pareto ranking is a fundamental selection method. Its effectiveness is demonstrated for a moderate number of objectives, which are subject to optimization simultaneously. With respect to the conflicting objectives in a MO optimization, one has to deal with the criteria as measures of the conflicts. The increased satisfaction of one criterion implies loss with respect to satisfaction of another criterion. Regarding to this, the formation of the Pareto front is based on some newly defined objective functions of the weighted N objectives $f_1, f_2, \ldots, f_N$ which are of the form

$$F_i(\boldsymbol{x}) = f_i(\boldsymbol{x}) + \sum_{j=1,\, j\neq i}^{j=N} a_{ji} f_j(\boldsymbol{x}), i = 1, 2, ..., N \tag{1}$$

where $F_i(x)$ are the new objective functions; a_{ji} is the designated amount of gain in the j-th objective function for a loss of one unit in the i-th objective function. Therefore the sign of a_{ji} is always negative. The above set of equations requires fixing the matrix $\boldsymbol{a}$. This matrix has all ones in its diagonal elements. To find the Pareto front of a maximization problem we assume that a solution parameter vector x_1 dominates another solution $\boldsymbol{x}_2$ if $F(\boldsymbol{x}_1) \geq F(\boldsymbol{x}_2)$ for all objectives. At the same time a contingent equality is not valid for at least one objective. $F_i(x)$ functions define the contour lines, which form a convex hull with the coordinate axes. This is illustrated in figure 1. In the case a_{ji} becomes zero then the contour lines are horizontal and vertical and $F_i(\boldsymbol{x}) = f_i(\boldsymbol{x})$. Explicitly, figure 1 shows the contour lines corresponding to two linear functions for a two-objective MO case. In this case the objectives are in particular subject to maximization. From the figure it is important to note that the point P is ultimately

$$\begin{aligned} F_1(\boldsymbol{x}) &= f_1(\boldsymbol{x}) \\ F_2(\boldsymbol{x}) &= f_2(\boldsymbol{x}) \end{aligned} . \tag{2}$$



(a) (b)

Fig. 1. Contour lines defining the dominated region via a reference point P (a); domains of relaxation that are hatched (b)

subject to identification as an ideal solution. The point P can be sought by means of appropriate methodologies one of which is the method of genetic algorithm applied in this approach. For the search process itself, different strategies can be followed. Below, two strategies are described for the sake of providing insight into the search process of the multi objective optimization approach applied in this work.

In one strategy the point P denotes the explicit ultimately attainable goal defined by the boundaries of the objective functions. The premise of the strategy is that beyond this point the solution is not acceptable or meaningless due to the limits of the objective functions.

The algorithm using the orthogonal lines is called *greedy* MO algorithm. The modification of the contour lines departing from horizontal and vertical ones defines a modified solution space. This is shown in figure 1b by hatched areas. These areas are termed as domain of relaxation in this work. Solutions found in these areas are not valid Pareto solutions, although they seemingly are solutions. For this modified solution space the area of non-existent solutions in the convex hull is diminished. This is at the expense of deviating from the strict non-dominated solution condition as shown in figure 1a. From figure 1b it should be noted that in case the strict non-domination condition is relaxed the Pareto front is smaller compared to the case of greedy search in figure 1a. However, this might be compensated by moving the Pareto surface forward more with pressure, so that the front comes closer to the point P. This strategy in the parameter space allows selection of the parameters in such a way that the relaxation domains in the objective functions space come to existence.

Figure 1b and its reference point serve as a conceptual explanation of the Pareto-optimality in order to point-out the 'trade-off' inherent to the relaxed dominance compared to the greedy dominance concept. Namely, with reference to the point P, by making the angle θ larger than 90 degrees the area of non-existent solutions is reduced compared to the greedy case. Therefore the Pareto front is allowed to establish closer to the reference point P, while at the same time the front is expected to be more diverse. In figure 1b it is seen how the greedy front comes closer to the point P through the widening of the angle θ. This is indicated by means of arrows. In the relaxed approach, the avoidance of aggregation to some extend is due to the distortion of the objective space, where the space becomes larger, and thus the density of solutions per unit length along the front is expected to become lower.

In a second strategy a hypothetical point designated as P′ denotes the explicit predefined sub-attainable goal. This goal is positioned somewhere in the convex hull defined via *P*, and is hopefully not far from the point *P*. This is shown in figure 2. It is to note that, since the point P is not explicitly known, the position assessment of the point P′ is a problematic issue. In any case P′ is a sub-optimal point implying some form of relaxation in the search process. In this case, the premise of the strategy is that the increased size of the area of non-existent solutions for orthogonal search domains is compensated with the increase of the size of the Pareto front in the relaxed case. At the same time the area of non-existent solutions is reduced for relaxed search domains. This is seen from the figure 2, where Pareto fronts having orthogonal and non-orthogonal contour lines are shown together. Domains of relaxations are also indicated in figure 2b. The latter strategy in the parameter space allows selecting the parameters in such a way that the relaxation domains in the objective functions space of figure 1b do not come to existence. However, P′ may come reasonably close to P, while this may not happen too. In the latter case still a Pareto front can be obtained. However the front remains to be poor. Due to this very reason, it is noteworthy to point out that in effective multi-objective optimization to obtain a Pareto surface is only half of the task to be accomplished. The other half is to obtain an effective front at the same time. The second strategy may not allow the parameters to explore the whole solution space due to the non-convex region defined by orthogonal and non-orthogonal axes in figure 2a. From the figure we note that as the degree of relaxation increases, the diversity of solutions along the Pareto surface increases, too. This is at the expense of reduced effectiveness of the multi objective optimization. Therefore, for effective optimization, the degree of relaxation in the search process should be kept to minimum being independent of the method of search. Although the Pareto front concept is well defined, the formation of this front is dependent on the implementation of the MO algorithm. Also it depends on the nature of the application. One interesting application is *the cone domination* approach (Branke et al., 2000), where the a_{ij} coefficients in (1) are determined in advance, so that the optimality search is carried out in the F-space in place of f-space. However, the necessity of adaptive modification of the coefficients during the search makes the approach rather impractical in the higher dimensions.

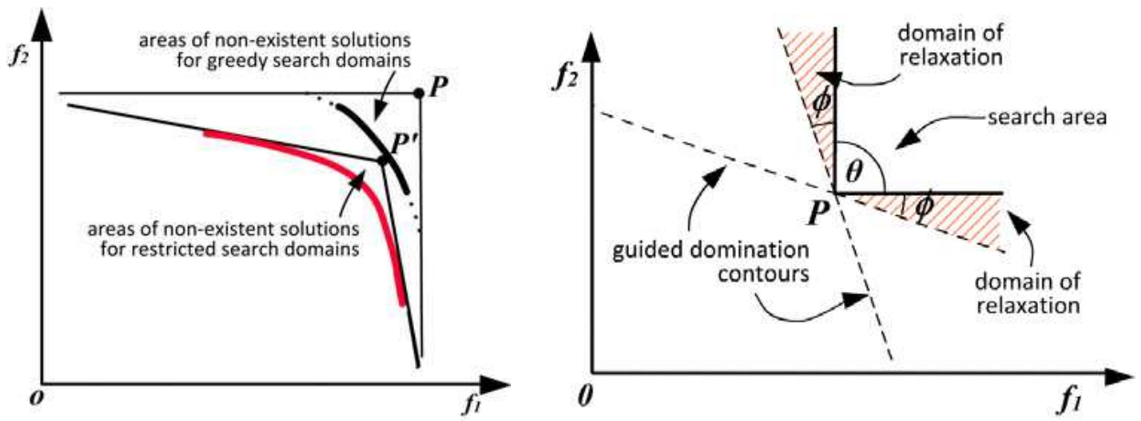


Fig. 2. Contour lines defining the dominated regions in orthogonal versus relaxed search

For the *greedy* application of the MO algorithm, one uses the orthogonal contour lines at the point P as shown in figure 2b. In this figure the point P denotes one of the individuals among the population in the context of genetic algorithm based evolutionary search. In the *greedy* search many potential favourable solutions are prematurely excluded from the search process. This is because during the search each solution of the population is represented by the point P, and the dominance is measured in relation to the number of solutions falling into the search domain within the angle $\theta=\pi/2$. This constant angle is the clear manifestation of the non-adaptivity of the conventional EMO algorithms.
Based on the strategy depicted in figure 2b the search process can be relaxed in a controlled way concerning the trade-off between solution diversity and effective multi-objective optimisation. That is the mechanism pushing the Pareto front forward is well balanced. This is a novel approach and it is explained below.

2.1 Relaxed Pareto ranking

To avoid premature elimination of potential solutions, a relaxed dominance concept is implemented, where the angle θ can be considered as the *angle of tolerance* provided $\theta>\pi/2$; the wider the angle beyond $\pi/2$, the more tolerant the search process, and vice versa. For $\theta<\pi/2$, the search becomes commensurately more greedy, and θ represents the *angle of greediness* in this case. In figure 2 the trade-offs between the *greedy* dominance and relaxed dominance methods are seen. In the earlier case the solutions are expectedly more effective due to their non-dominance, but diverse solutions are quickly exhausted due to non-adaptivity, and therefore the solutions are aggregated. In the latter case, the solutions are more diversified but preliminarily less effective. However, in the long run, the diverse solutions can develop to more effective solutions compared to those obtained from the greedy dominance approach. The implementation of the relaxed dominance approach may be simple in two-dimensional objective space as the hatched relaxation domain is easy to deal with as seen in figure 2b. However, in multi-dimensional objective space, the same domain goes even beyond a simple imagination. To be able to deal with the relaxed domain a mathematical method is developed in this work which is described below.
Let (1) be expressed by

$$\begin{aligned} F_1 &= f_1 \quad + a_{21} f_2 + \ldots\ldots + a_{n1} f_n \\ F_2 &= a_{12} f_1 + f_2 \quad + \ldots\ldots + a_{n2} f_n \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ F_n &= a_{1n} f_1 + a_{2n} f_2 + \ldots\ldots + a_{nn} f_n \end{aligned} \quad . \tag{3}$$

In matrix equation form, (3) becomes

$$F = \begin{bmatrix} F_1 \\ F_2 \\ \ldots \\ F_n \end{bmatrix} = \begin{bmatrix} a_{11} & a_{21} & \ldots\ldots & a_{n1} \\ a_{12} & a_{22} & \ldots\ldots & a_{1n} \\ \ldots\ldots\ldots\ldots & & & \\ a_{1n} & a_{2n} & \ldots\ldots & a_{nn} \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \\ \ldots\ldots \\ f_n \end{bmatrix} = \begin{bmatrix} 1 & a_{21} & \ldots\ldots & a_{n1} \\ a_{12} & 1 & \ldots\ldots & a_{1n} \\ \ldots\ldots\ldots\ldots & & & \\ a_{1n} & a_{2n} & \ldots\ldots & 1 \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \\ \ldots\ldots \\ f_n \end{bmatrix} \tag{4}$$

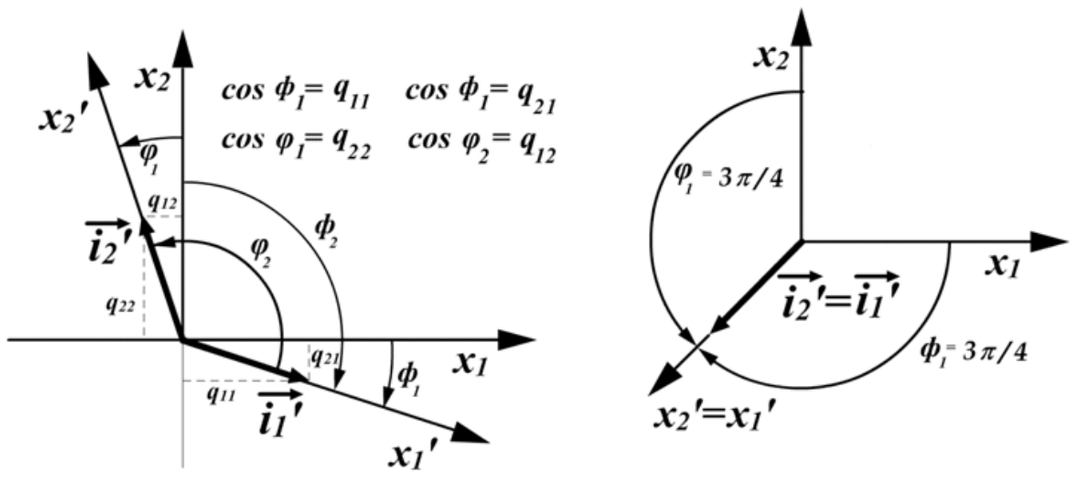
where $a_{11}, a_{22}, \ldots, a_{nn}$ are equal to unity. For the sake of simplicity in the description below only two objectives are considered while the results are valid for any dimension. The objective functions for this case are given by

$$F_1 = f_1 + a_{21}\, f_2$$
$$F_2 = a_{12}\, f_1 + f_2 \tag{5}$$

In a two-dimensional coordinate system the contour lines in figures 1a, 1b and 2b are orthogonal and non-orthogonal respectively. The search area in the latter case includes also the domains of relaxation. These are added to the search area of the orthogonal system, as illustrated in figure 2b.

In the non-orthogonal system the search area for the favourable solutions is wider. At the same time some of the solutions are not dominating the solution at the point P seen in figure 2b. However, as a trade-off it provides more diversity at the final Pareto front, while the front is not entirely non-dominated. The solutions at the front are more probably non-dominated in the middle part of the front. This is where f_1 and f_2 are close to each other. Conversely, the solutions may be more dominated at the regions close to edges of the front (Deb et al., 2003). This is a property of the *cone dominance* approach. This situation occurs since in the cone domination approach a greedy algorithm is applied using a non-orthogonal system taking a reference point P as origin. By doing so, the search algorithm remains the same but it uses the coordinates of the new non-orthogonal system. However, this cone dominance approach does not address the problem of aggregation. This becomes especially problematic in higher multi-dimensional optimization. This means, the Pareto front is potentially wider in the F-space. However, without resolving the aggregation phenomenon the potentially wider Pareto front remains ineffective. In contrast to the cone dominance approach, in this work each member of the population is considered to be represented by point P seen in figure 2b, and the solutions falling into the relaxation domains are included to the non-dominated solutions, thereby accruing some dominated solutions to the non-dominated ones to form the next-generation solutions. This means, the orthogonal system is not replaced by the non-orthogonal system but the greedy non-dominated orthogonal search space is relaxed. The relaxed domains simply contribute to the greedy search domain with some additional, potentially lucrative solutions.

Interestingly, this situation is similar to the classical gradient-based optimization method where each iteration the step-length towards the global maxima or minima determined by a step-size parameter (Farhang-Boroujeny, 1998), also called convergence coefficient. The step-size parameter should be small enough to ensure the stability of the convergence (Bazaraa et al., 1993, Kuester, 1973). If the step-size parameter is zero, the approach to minima or maxima does not occur. If it is too big, convergence does not occur, due to instability. For similar reasons, in the evolutionary computation the angle ϕ defining the relaxation domain should be kept small. In this way the stability of the algorithm is maintained and the effectiveness is enhanced. It should be noted that, although the angle ϕ is small it plays role for each population member at each generation. This makes the net effect of the relaxation highly significant. In the adaptive Pareto front formation $\cos(\phi)$ plays the role of convergence coefficient defined in the gradient-based optimization. If $\cos(\phi)$ is maximum, i.e. $\cos(\phi)=1$, greedy search in orthogonal system occurs, leading to aggregation. This can be considered as a kind of instability in this case. From the other side, if $\cos(\phi)$ is minimal, *i.e.* $\cos(\phi)=-1$, convergence does not occur, because the evolutionary search process becomes trivial in this case. This situation already occurs for $\phi<-3\pi/4$, which corresponds to $\cos(\phi)=-.707$. This is illustrated in figure 3b.



(a) (b)

Fig. 3. Direction cosines in a coordinate transformation in 2-dimensional case (a); schematic representation of relaxation domains covering the total parameter search space as an extreme situation (b)

In (4) the small-enough designation of the parameters a_{ji} is crucial for the performance of the evolutionary computation. They characterize the cosines of the angle between the respective coordinate axes. In general the angle ϕ is application dependent. The normalization of these cosines yields the directive cosines, which determine the transformation from orthogonal to the non-orthogonal, i.e., relaxed coordinate system. If we denote the direction cosines as q_{ij}, the transformation matrix becomes

$$Q = \begin{bmatrix} q_{11} & q_{12} & \ldots\ldots & q_{1n} \\ q_{21} & q_{22} & \ldots\ldots & q_{2n} \\ \ldots\ldots\ldots\ldots & & & \\ q_{n1} & q_{n2} & \ldots\ldots & q_{nn} \end{bmatrix}, \tag{6}$$

which transforms the non-orthogonal system to the orthogonal system and vice versa via

$$x = Q\,x' \tag{7}$$

$$x' = Q^{-1}\,x \tag{8}$$

where x' denotes the non-orthogonal system $x'=[x_1', x_2', \ldots, x_n']^T$. The directive cosine row vectors of (4) are given by

$$d_i = \left[d_{1i}\; d_{2i}\; .. d_{ni}\right] = \frac{1}{\sqrt{\sum_{j=1}^{n} a_{ji}^2}} [a_{1i}\;\; a_{2i}\;\; \ldots\; a_{ni}] \tag{9}$$

which corresponds to column vectors in (6), so that

$$D = \begin{bmatrix} [d_{11}\ d_{21}\ .....\ d_{n1}] \\ [d_{12}\ d_{22}\ .....\ d_{n2}] \\ ..... \\ [d_{1n}\ d_{2n}\ .....\ d_{nn}] \end{bmatrix} = Q^T \tag{10}$$

and

$$\sqrt{\sum_{i=1}^{n} q_{ji}^2} = 1 \qquad j = 1, 2, ..., n \tag{11}$$

In a two-dimensional case the directive cosines are shown in figure 3a and in relation to the set of transformation equations in (3), the directive cosine row vectors are given by

$$\begin{aligned} d_1 &= \begin{bmatrix} d_{11}\ d_{21} \end{bmatrix} = \begin{bmatrix} 1/\sqrt{1+a_{21}^2} & a_{21}/\sqrt{1+a_{21}^2} \end{bmatrix} \\ d_2 &= \begin{bmatrix} d_{12}\ d_{22} \end{bmatrix} = \begin{bmatrix} a_{12}/\sqrt{1+a_{12}^2} & 1/\sqrt{1+a_{12}^2} \end{bmatrix} \end{aligned} \tag{12}$$

The coordinate transformation of point *P* in figure 2b is given by

$$\begin{bmatrix} f_1^P \\ f_2^P \end{bmatrix} = \begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix} \begin{bmatrix} F_1^P \\ F_2^P \end{bmatrix} = \begin{bmatrix} d_{11} & d_{21} \\ d_{12} & d_{22} \end{bmatrix} \begin{bmatrix} F_1^P \\ F_2^P \end{bmatrix} \tag{13}$$

and

$$\begin{bmatrix} F_1^P \\ F_2^P \end{bmatrix} = \begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix}^{-1} \begin{bmatrix} f_1^P \\ f_2^P \end{bmatrix} = \begin{bmatrix} d_{11} & d_{21} \\ d_{12} & d_{22} \end{bmatrix}^{-1} \begin{bmatrix} f_1^P \\ f_2^P \end{bmatrix} \tag{14}$$

It is interesting to note that since the cosine directive q_{11} in (13) given by $q_{11}=cos(\phi_1)$ is equal to d_{11} in (12), so that

$$q_{11} = \cos(\phi_1) = 1/\sqrt{1+a_{21}^2}$$

and using the relationship

$$\cos(z) = 1/\sqrt{1+\tan(z^2)} \tag{15}$$

we obtain

$$a_{21} = \mp \tan(\phi_1) \tag{16}$$

which yields (5) in terms of the transformation angles in the form

$$F = \begin{bmatrix} F_1 \\ F_n \end{bmatrix} = \begin{bmatrix} 1 & a_{21} \\ a_{12} & 1 \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} = \begin{bmatrix} 1 & \tan(\phi_2) \\ \tan(\varphi_2) & 1 \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} \tag{17}$$

In a general form, (17) is given by

$$F=\begin{bmatrix}F_1\\F_2\\...\\F_n\end{bmatrix}=\begin{bmatrix}1\ a_{21}\ .....\ a_{n1}\\a_{12}\ 1\ .....\ a_{1n}\\.....\\a_{1n}\ a_{2n}\ .....\ 1\end{bmatrix}\begin{bmatrix}f_1\\f_2\\.....\\f_n\end{bmatrix}=\begin{bmatrix}1 & \tan(\phi_2) & ..... & \tan(\phi_n)\\ \tan(\varphi_2) & 1 & ..... & \tan(\varphi_n)\\ ..... & & & \\ \tan(\theta_2) & \tan(\theta_n) & ..... & 1\end{bmatrix}\begin{bmatrix}f_1\\f_2\\.....\\f_n\end{bmatrix} \tag{18}$$

The importance of the coordinate transformation becomes dramatic especially in higher dimensions. In such cases the spatial distribution of domains of relaxation becomes complex and thereby difficult to implement. Namely, in multidimensional space the volume of a relaxation domain is difficult to imagine. And more importantly it is difficult to identify the population in such domains. Therefore one needs a systematic approach for identification by computation and not by inspection or anything else. This systematic approach is based on the coordinate transformation as follows. Basically for each solution point, designated in general as P in figure 2b is temporarily considered to be a reference point as origin and all the other solution points in the orthogonal coordinate system are converted to the non-orthogonal system coordinate by (8). For instance for four objectives, we write

$$F=\begin{bmatrix}F_1\\F_2\\F_3\\F_4\end{bmatrix}=\begin{bmatrix}1\ a_{21}\ a_{31}\ a_{41}\\a_{12}\ 1\ a_{32}\ a_{42}\\a_{13}\ a_{23}\ 1\ a_{43}\\a_{14}\ a_{24}\ a_{34}\ 1\end{bmatrix}\begin{bmatrix}f_1\\f_2\\f_3\\f_4\end{bmatrix}, \tag{19}$$

where the a_{ij} parameters determine the relaxation angles and consequently the cosine directives that are computed from the relaxation angle, that is modified during the search adaptively. The relaxation of the dominance with the adaptive process guarantees the prevention of aggregation. As a numerical example, for a relaxation angle of $\phi=\varphi=\psi=\theta=7^o$ and symmetry similar to that seen in figure 3 of two-dimensional case, one obtains

$$\begin{bmatrix}F_1\\F_2\\F_3\\F_4\end{bmatrix}=\begin{bmatrix}1 & -.123 & -.123 & -.123\\ -.123 & 1 & -.123 & -.123\\ -.123 & -.123 & 1 & -.123\\ -.123 & -.123 & -.123 & 1\end{bmatrix}\begin{bmatrix}f_1\\f_2\\f_3\\f_4\end{bmatrix} \tag{20}$$

so that, in view of (7) the cosine direction matrix and the corresponding weighted objectives F_1 and F_2 are given by

$$\begin{bmatrix}f_1^P\\f_2^P\\f_3^P\\f_4^P\end{bmatrix}=\begin{bmatrix}.993 & -.122 & -.122 & -.122\\ -.122 & .993 & -.122 & -.122\\ -.122 & -.122 & .993 & -.122\\ -.122 & -.122 & -.122 & .993\end{bmatrix}\begin{bmatrix}F_1^P\\F_2^P\\F_3^P\\F_4^P\end{bmatrix} \tag{21}$$

And in view of (8) the inverse of (21) becomes

$$\begin{bmatrix} F_1^P \\ F_2^P \\ F_3^P \\ F_4^P \end{bmatrix} = \begin{bmatrix} 1.072 & .174 & .174 & .174 \\ .174 & 1.072 & .174 & .174 \\ .174 & .174 & 1.072 & .174 \\ .174 & .174 & .174 & 1.072 \end{bmatrix} \begin{bmatrix} f_1^P \\ f_2^P \\ f_3^P \\ f_4^P \end{bmatrix} \tag{22}$$

After conversion, all points which have positive coordinates in the non-orthogonal system correspond to potential solutions contributing to the next generation in the evolutionary computation. If any point possesses a negative component in the new coordinate system, the respective solution does not dominate P and therefore is not counted. This is because otherwise such a solution may lead the search in a direction away from P. In general, the relaxation of the dominance in higher dimensions is extremely complex and therefore many different methods for effective Pareto front formation in the literature (Hughes, 2005, Jaszkiewicz, 2004) are reported. However (8) provides a decisive and easy technique revealed in this work for the same goal.

3. Adaptive formation of Pareto front in a design application

In this section adaptive formation of Pareto front with multi-objectivity is considered. The formation of Pareto front is explained by means of a design example where multi-objectivity is subjected to a Pareto front based solution. For the multi-objective optimization a genetic algorithm approach with a relaxed Pareto-ranking is used. The relaxation angle is computed adaptively for every chromosome, and at every generation. This is implemented by having the angle be a part of the chromosome of every solution. The fitness of a chromosome is obtained by considering two properties of the solution at the same time. One is the degree of dominance in terms of the amount of solutions dominating an individual, the second is the relaxation angle used to measure this amount. This is given by

$$R_{fit} = \frac{1}{N(\theta) + n} \tag{23}$$

$$N(\theta) = \frac{4}{1 + \left(\theta / \bar{\theta}\right)} \tag{24}$$

In (23) and (24) the purpose is to reward a chromosome for affording a wide relaxation angle θ, relative to the average angle of the population $\bar{\theta}$, and still having a low dominance count, denoted by n. The wide angle provides more diversity in the population for the next generation. However, when relaxation angle would be excessively big, the population for the next generation can be crowded with trivial solutions. To prevent that, in (23) the number of non-dominated solutions with respect to the particular solution considered denoted by n, is summed up with the function of the angle $N(\theta)$. This means that between two solutions with the same amount of non-dominated solutions, the one with the wider angle is preferred. This is done for every solution in the population. This implies that the average angle $\bar{\theta}$ is changing for every generation adaptively. It is noted that the $N(\theta)$ appearing in (24) is, used to adjust the relative importance of relaxation angle versus count

n. Figure 4 shows the mean relaxation angle during the adaptive Pareto front formation by the genetic algorithm. The angle converges to a fixed angle of 7^o as seen from the figure. This is a clear indication of the stochastic adaptive process.

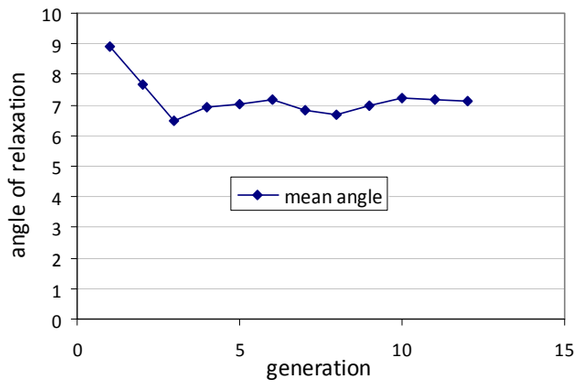


Fig. 4. Variation of mean relaxation angle during the adaptive Pareto front formation by the GA

Since design is an intelligent activity it is composed of several considerations which are soft in nature. For this reason, to invoke methods of soft computing is most appropriate. In this respect fuzzy-neural tree is considered as a knowledge model in this work as it is explained below. The application concerns the design of a building. The building consists of a number of spatial units, referred to as design objects, where every unit is designated to a particular purpose in the building. The task is to locate the objects optimally with respect to three objectives. The objects are seen from figure 5.

Object O1 is a hotel unit, O3 is a conference space, O4 and O5 are office wings, and O2 is a lobby with a restaurant. In order to let the computer generate a building from these parts, i.e. for the solution to be feasible, it is necessary to ensure that all solutions have some basic properties. These are that spaces should not overlap, and objects should be adjacent to the other objects around. This is realized in the present application by inserting the objects in a particular sequential manner into the site. This is illustrated in figure 5. One by one the objects are moved into the site starting from a point marked by a cross in the figures, then moving in west direction (lower part of the figures) until they reach an obstacle, that may be the site boundary or another object previously inserted. When they touch an object they change their movement direction from western to the northern direction, moving north until they again reach the site boundary or another object. As a third and final movement step the object will attempt to move once more in western direction, although often this is may not happen since often there is an adjacent object in the western direction blocking the way. An example of this third step is shown in figure 5c. The third step is to avoid that gaps between objects are reduced to some extend. This way of packing objects is known as *bottom-left* method in literature, and it is used mainly for packing problems. After the final object has been placed in this way the design is ready for evaluation. It is noted that the decision from which side to insert the objects, and which location to use for the insertion point is a matter of judgement, and it will strongly influence the solutions obtained. The insertion used in this application is due to the preference of the architect is to have the objects line up along the street, which is in western boundary of the site.

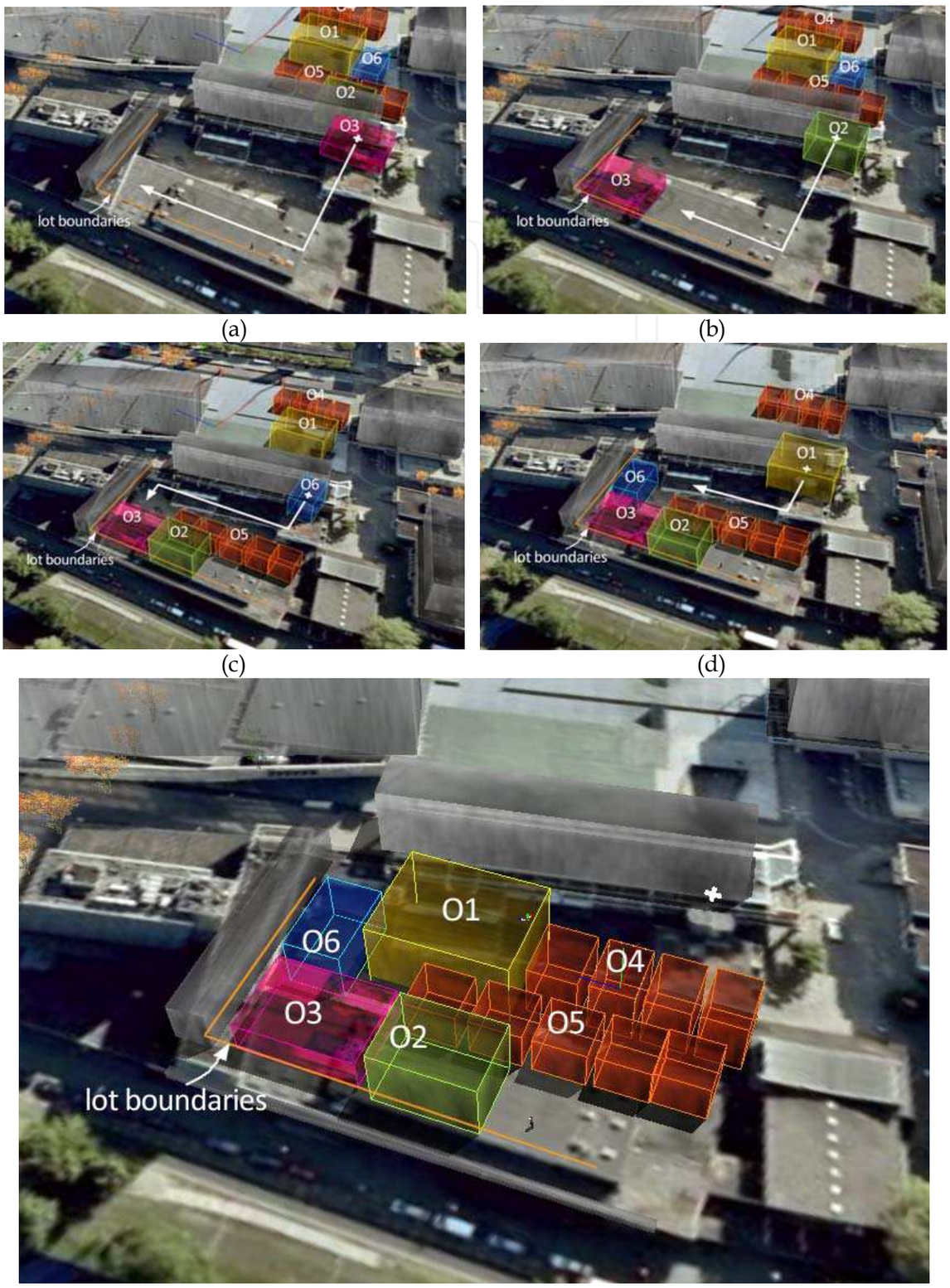


Fig. 5. Generation process of a solution through sequential insertion of the design objects

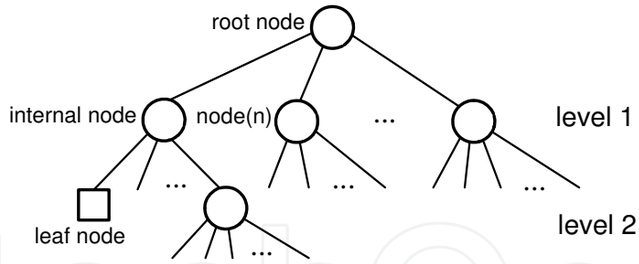


Fig. 6. The structure of a neural tree

The multi objective optimization is accomplished using a multi-objective genetic algorithm with a relaxed Pareto ranking. It is used to determine the optimal sequence of insertion, so that three objectives are maximally fulfilled. Objects are numbered, and every chromosome contains the information for every object, at which rank in the insertion sequence it is to be inserted.

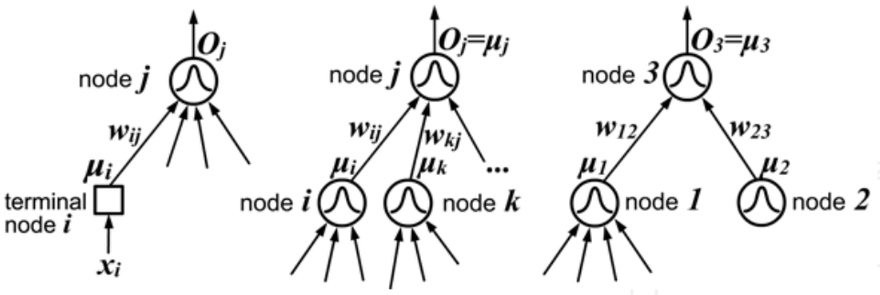
The objectives are functionality of the building, certain energy performance aspects, and some form related preferences. The design performance is obtained from these three objectives. Due to the linguistic nature of these objectives a special model is formed and used to assess satisfaction of the objectives. This model is a fuzzy neural tree model. In this model the ultimate goal is to have a good design, what we term as a design with a high design performance. That is, if all three objectives are highly fulfilled then the design has a high performance. The relation of the concept of design performance with the physical properties of possible solutions, which form the model inputs, is captured in the model through a hierarchical structure of logic operations. The method used is fuzzy neural tree.

3.1 Fuzzy-neural tree modeling domain knowledge

For human-like information processing the methods of soft computing are presumably the most convenient. The salient soft computing methods are in the paradigms of neural nets and fuzzy logic (Mitra et al., 2002). In this work a neural tree is considered to assess the suitability of a solution in a human-like manner. A neural tree is composed of terminal nodes, non-terminal nodes, and weights of connection links between two nodes. The non-terminal nodes represent neural units and the neuron type is an attribute introducing a non-linearity simulating a neuronal activity. In the present case, this attribute is established by means of a Gaussian function which has several desirable features for the intended goals; namely, it is a radial basis function ensuring a solution and the smoothness. At the same time it plays the role of a fuzzy membership function in the tree structure, which is considered to be a fuzzy logic system as its outcome is based on fuzzy logic operations and thereby associated reasoning. An instance of a neural tree is shown in figure 6.

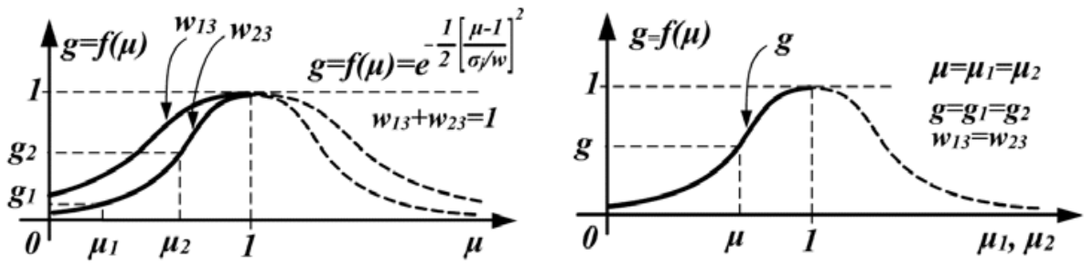
Detailed structures of a neural tree are shown in figure 7. Figure 7a shows a terminal node connected to an inner node, and figure 8b and 8c show the connections among inner nodes. Each terminal node, also called *leaf*, is labelled with an element from the terminal set $T=\{x_1, x_2, \ldots, x_n\}$, where x_i is the i-th component of the external input vector $\boldsymbol{x}$. Each link (i,j) represents a directed connection from node i to node j. A value w_{ij} is associated with each link as seen from figure 7. In a neural tree, the root node is an output unit and the leaf nodes, or terminal nodes, are input units. The node outputs are computed in the same way as computed in a feed-forward neural network. In this way, neural trees can represent a broad class of feed-forward networks that have irregular connectivity and non-strictly

layered structures. In particular, in the present work the nodes are similar to those used in a radial basis functions network with the Gaussian basis functions.



(a) (b) (c)

Fig. 7. Detailed structures of a neural tree with respect to different type of node connections



(a) (b)

Fig. 8. Fuzzification for two inputs (a), fuzzification where $\mu = \mu_1 = \mu_2$ (b)

In the neural tree considered in this work the output of i-th node is denoted x_i and it is introduced to another node j. A non-terminal node consists of a Gaussian radial basis function.

$$f(X) = w\phi(\| X - c \|^2) \tag{25}$$

where $\phi(.)$ is the Gaussian basis function, $\boldsymbol{c}$ is the center of the basis function. The Gaussian is of particular interest and used in this research due to its relevance to fuzzy-logic. The width of the basis function σ_j at node j is used to measure the uncertainty associated with the inputs to this node, designated as external input X_j. X_j is related to the output of node i denoted as μ_i by relation

$$X_j = \mu_i w_{ij} \tag{26}$$

where w_{ij} is the weight connecting node i to node j. The centers of the basis functions are the same as the input weights of that node.
The output of node j is given by

$$O_j = \exp(-\frac{1}{2}\sum_i^n \left[\frac{w_{ij}\mu_i - w_{ij}}{\sigma_j}\right]^2) \tag{27}$$

which reduces to

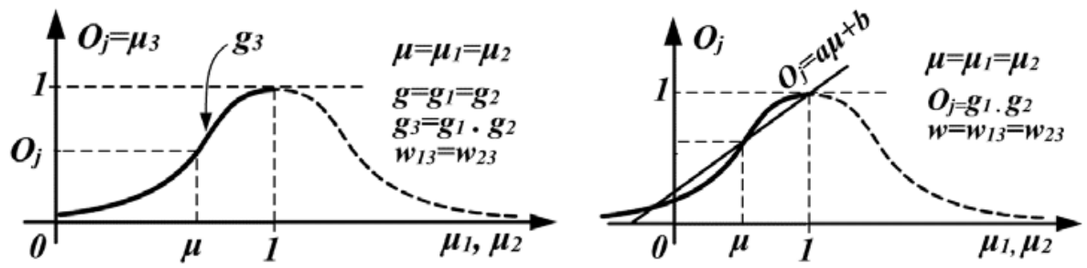
$$O_j = \exp(-\frac{1}{2}\sum_i^n \left[\frac{w_{ij}(\mu_i - 1)}{\sigma_j}\right]^2) \tag{28}$$

We can express (28) in the following form

$$O_j = \exp(-\frac{1}{2}\sum_i^n \left[\frac{(\mu_i - 1)}{\sigma_j / w_{ij}}\right]^2) \tag{29}$$

This implies that the width of the Gaussian is scaled by the input weight w_{ij}. In other words, as to the width, the shape of Gaussian fuzzy membership function is dependent on the input weights w_{ij} determined by the domain knowledge. It should be noted that this is a novel type of computation at each node which is quite different than conventional radial basis function (RBF) type computation, where the centers are determined by other means, clustering for instance. However, a non-terminal node itself can be seen as an RBF having different width for each dimension. For such a node, there should be at least two inputs with appropriate connection weights. The connection weights of a node should be normalized, so that the sum of the weights becomes equal to 1.

In figure 8 only two inputs are considered without loss of generality. The variables w_{13} and w_{23} are input weights determining the width of the Gaussian membership functions. It is noted that in the figure $w_{13}<w_{23}$. This is clear from (29). For two inputs, two distinct standard deviations are defined. In particular, the inputs can be equal, i.e., $\mu_1 = \mu_2$. This particular case occurs when the outputs of the two nodes delivering the inputs to the node we are considering are equal. This case is illustrated in figure 9b. In figure 9a it is clear that, if w_1 and w_2 are equal then the AND operation is expressed by means of a single Gaussian denoted by g_3.



(a) (b)

Fig. 9. And operation (a); linear approximation to Gaussian function (b)

Since σ_j is a free parameter, by giving an appropriate width via (29), the result of the AND operation is given by

$$\begin{aligned} &\mu = \mu_1 = \mu_2 \\ &f(\mu) = f(\mu_1) = f(\mu_2) = g_1 = g_2 \\ &O_j = g_1 \cdot g_2 \end{aligned} \tag{30}$$

This is illustrated in figure 9b where the left part of the Gaussian is approximated by a straight line. In figure 9b, optimizing the σ_j parameter, we obtain

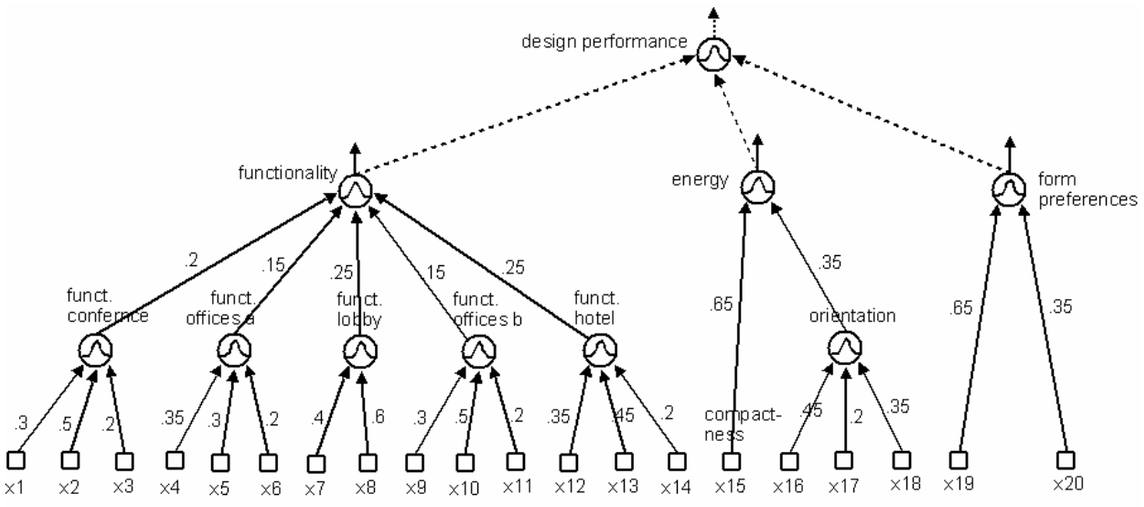


Fig. 10. Neural tree used to obtain fitness regarding three objectives

$$O_j \cong \mu \tag{31}$$

for the values μ and O_j that can take between zero and one. In any case, for a node in the neural tree, (31) is satisfied for $\mu=O_j=0$ (approximately) and for $\mu=O_j=1$ (exact) inherently, while g_1 and g_2 are increasing function of μ_1 and μ_2. Therefore a linear relationship between O_j and μ in the range between 0 and 1 is a first choice from the fuzzy logic viewpoint; namely, as to the AND operation at the respective node, if inputs are equal, that is $\mu=\mu_1=\mu_2$ then the output of the node of μ_1 AND μ_2 is determined by the respective *triangular membership functions* in the antecedent space. Triangular fuzzy membership functions are the most prominent type of membership functions in fuzzy logic applications. For five inputs to a neural tree node, these membership functions are represented by the data sets given by Table 1 and Table 2

.1	.2	.3	.4	.5	.6	.7	.8	.9
.1	.2	.3	.4	.5	.6	.7	.8	.9
.1	.2	.3	.4	.5	.6	.7	.8	.9
.1	.2	.3	.4	.5	.6	.7	.8	.9
.1	.2	.3	.4	.5	.6	.7	.8	.9

Table 1. Data-set at neural tree node input

.1	.2	.3	.4	.5	.6	.7	.8	.9

Table 2. Data-set at neural tree node input

In general, the data sets given in Table 1 and Table 2 are named in this work as '*consistency conditions*'. They are used to calibrate the membership function parameter σ. This is accomplished by optimization.

At this point a few observations are due, as follows. If a weight w_{ij} is zero, this means the significance of the input is zero, consequently the associated input has no effect on the node output and thus also the system output. Conversely, if a w_{ij} is close to unity, this means the significance of the input is highest among the competitive weights directed to the same

node. This means the value of the associated input is extremely important and a small change about this value has big impact on the node output O_j. If a weight w_{ij} is somewhere between zero and one, then the associated input value has some possible effect on the node output determined by the respective AND operation via (29). In this way, the domain knowledge is integrated into the logic operations.

The general properties of the present neural tree structure are as follows.

If an input of a node is small (i.e., close to zero) and the weight w_{ij} is high, then, the output of the node is also small complying with the AND operation.

If a weight w_{ij} is low the associated input cannot have significant effect on the node output. This means, quite naturally, such inputs can be ignored.

If all input values coming to a node are high (i.e., close to unity), the output of the node is also high complying with the AND operation.

If a weight w_{ij} is high the associated input x_i can have significant effect on the node output.

It might be of value to point out that, the AND operation in a neural-tree node is executed in fuzzy logic terms and the associated connection weights play an important role on the effectiveness of this operation.

The neural tree employed in this work is shown in figure 10. The root node describes the ultimate goal subject to maximization, namely the design performance and the tree branches form the objectives constituting this goal. The connections among the nodes have a weight associated with them, as seen from the figure. The weight is given by a designer, as an expression of knowledge, and it specifies the relative significance a node has for the node one level closer to the root node. It is noted that in the multi-objective optimization case the weights connecting the nodes on the penultimate level of the weight tells how strongly the output of the lower node influences the output of the upper tree to the root node are not specified a-priori, but they are subject to identification after the optimization process is accomplished.

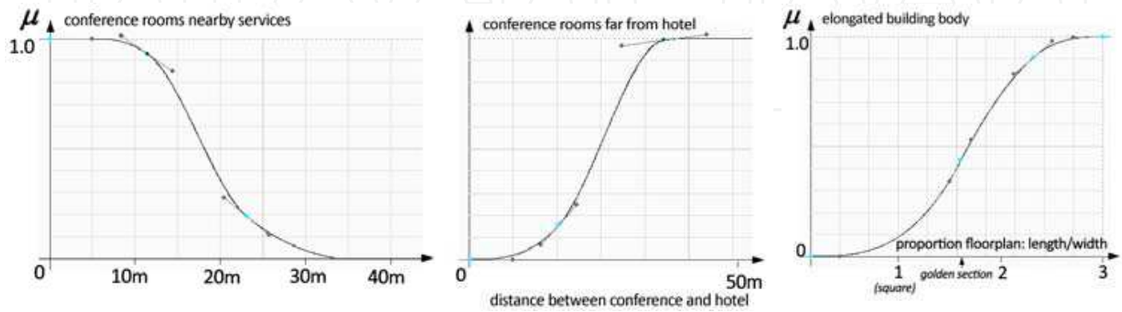
During evaluation of a design alternative the tree is provided with inputs at its leaf nodes and the fuzzification processes are carried out. The fuzzification yields the satisfaction of an elemental requirement at the terminal nodes of the neural tree. These requirements are some desirable features expressed by means of fuzzy membership functions at the terminal nodes of the tree. Three examples are shown in figure 11.

Figure 11a expresses the requirement x1 in figure 10. It demands that the conference space should be located close to the building services, which is an aspect of functionality, i.e. distance from service facilities should be low for convenient access of services. The requirement is fully satisfied, i.e. the membership degree μ becomes unity, when the distance among the objects is less than about 10m. Since distances are measured from geometric centres of both objects, this distance belongs to the case that both objects are adjacent. Distances beyond 30m mean that the conference room is not considered as being nearby services, so that this requirement is not fulfilled, i.e. μ becomes zero.

Figure 11b expresses the requirement x3. It demands that the conference space should be far from the hotel, to avoid acoustic disturbance and to keep people flows separate. From the figure we see that distances beyond 40m are considered to fully satisfy this demand. From figure 10 we note that, comparatively the requirement in figure 11a is considered 1.5 times more important than figure 11b regarding the functionality of the building.

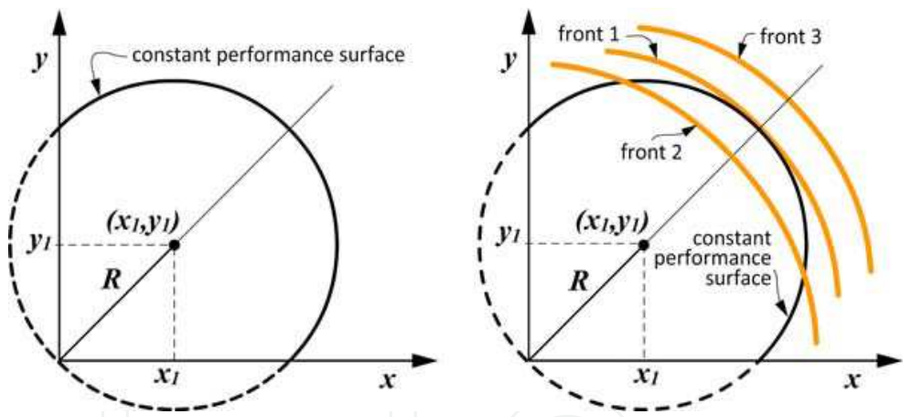
Figure 11c expresses the requirement x19 in figure 10. It demands that the building should have an elongated shape. What is meant is that the shape of the floor plan should not be a

square, but that the shape should clearly have a longer extend in north-south direction, termed length, than in east-west direction, termed width. This is an aspect of form preferences. To express this demand the input values to this membership function are the proportion length/width, being unit less. From the membership function we note that a square proportion yields a low membership degree μ, i.e. low satisfaction of this requirement. The particular proportion known as *golden section* yields approximately a satisfaction degree of μ=0.5 and proportion values beyond 2.5 are considered fully satisfying the requirement.



(a) (b) (c)

Fig. 11. Some membership functions at the terminal nodes of the neural tree in figure 10



(a) (b)

Fig. 12. Constant performance surface (a); analysis of three Pareto fronts (b)

In the same way other properties of the design are measured and converted into satisfactions using specific membership functions at the terminals. The fuzzified information is then processed by the inner nodes of the tree. These nodes perform the AND operations using Gaussian membership functions as described above. Finally this sequence of logic operations starting from the model input yield, the performance at the penultimate node outputs of the model. This means the more satisfied the elemental requirements at the terminal level are, the higher the outputs will be at the nodes above, finally increasing the design performance at the root node of the tree. Next to the evaluation of the design performance score, due to the fuzzy logic operations at the inner nodes of the tree, the performance of any sub-aspect is obtained as well. This is a desirable feature in design, which is referred to as transparency.

Having established the performance evaluation model, it is used for the evolutionary search process aiming to identify designs with maximal design performance. In the present case we are interested in a variety of alternative solutions that are equivalent in Pareto sense. The design is therefore treated as a multi-objective optimization as opposed to a single-objective optimization. In single-objective case exclusively the design performance, i.e. the output at the root node of the neural tree, would be subject to maximization. In the latter case, the solution would be the outcome of a mere convergence and any cognition aspect would not be exercised. In the multi-objective implementation the outputs of the nodes *functionality*, *energy*, and *form preferences*, which are the penultimate nodes, are subject to maximization. Their values are used in the fitness determination procedure of the genetic algorithm. Employing the fuzzy neural tree in this way the genetic search is equipped with some human-like reasoning capabilities during the search. The part of the tree beyond the penultimate nodes is for the defuzzification process, which models cognition, so that ultimately the design performance is obtained at the root node.

3.2 Design performance and the Pareto front

It is noted that generally multi-objective optimization involves no information on the relative importance among the objectives. Therefore, generally, Pareto optimal solutions cannot be distinguished without bringing into play other, i.e. higher-order criteria than the objectives used in the search. However, it is noted that the Pareto solutions may be distinguished as follows. From figure 10, at the root node, the performance score is computed by the defuzzification process given by

$$w_1 f_1 + w_2 f_2 + w_3 f_3 = p \tag{32}$$

where f_1 is the output of the node *functionality;* f_2 of node *energy*; f_3 of node *form preferences.* That is, they denote the performance values for these aspects of the design, which are subject to maximization. The variable p denotes the design performance which is also requested to be maximized. In (32) w_1 , w_2 , and w_3 denote the weights associated to the connections from *functionality*, *energy*, and *form preferences.* It is noted that $w_1+w_2+w_3=1$.
In this design exercise, the cognitive design viewpoint plays important role. This means it is initially uncertain what values $w_1,\ldots w_3$ should have. Namely, the node outputs $f_{1,\ldots,}f_3$ can be considered as the *design feature vector,* and the reflection of these features can be best performed if the weights w_1 ; …; w_3 define the same direction as that of the feature vector. Hence the components of the unit vector along the feature vector are computed as

$$u_1 = \frac{f_1}{\sqrt{f_1^2 + f_2^2 + \ldots + f_n^2}}; \quad u_2 = \frac{f_2}{\sqrt{f_1^2 + f_2^2 + \ldots + f_n^2}}; \ \ldots; \ u_n = \frac{f_n}{\sqrt{f_1^2 + f_2^2 + \ldots + f_n^2}} \tag{33}$$

Normalising the components and equating them to the weights yields

$$w_1 = \frac{f_1}{f_1 + f_2 + f_3}; \quad w_2 = \frac{f_2}{f_1 + f_2 + f_3}; \quad w_3 = \frac{f_3}{f_1 + f_2 + f_3} \tag{34}$$

In general, if there are n objectives at the penultimate layer of the neural tree, we can write that

$$u_1 = \frac{f_1}{\sqrt{f_1^2 + f_2^2 + ... + f_n^2}} \; ; \ldots ; \quad u_n = \frac{f_n}{\sqrt{f_1^2 + f_2^2 + ... + f_n^2}} \tag{35}$$

Above computation implies that, the performance p for each genetic solution is given by

$$p = \frac{f_1^2 + f_2^2 + ... + f_n^2}{f_1 + f_2 + ... + f_n} \tag{36}$$

Therefore, (36) is computed for all the design solutions on the Pareto front. Then the *solution with maximal performance* is selected among the Pareto solutions. This way the particular design is identified as a solution candidate with the corresponding w_1, w_2, …., w_n weights. These weights form a priority vector $\boldsymbol{w}^*$. In the present application (36) becomes

$$p = \frac{f_1^2 + f_2^2 + f_3^2}{f_1 + f_2 + f_3} \tag{37}$$

where f_n n-th output on the penultimate level of the neural tree. If for any reason this candidate solution is not appealing, the next candidate is searched among the available design solutions with a desired design feature vector and the relational attributes, i.e., w_1, w_2, …., w_n . One should note that, although performance does not play role in the genetic optimization, Pareto front offers a number of design options with fair performance leaving the final choice dependent on other environmental preferences. Using (36), second-order preferences are identified that are most promising for the task at hand, where ultimately maximal design performance is pursued.

To this end, to make the analysis explicit we consider a two-dimensional objective space. In this case, (36) becomes

$$p = \frac{f_1^2 + f_2^2}{f_1 + f_2} \tag{38}$$

which can be put into the form

$$f_1^2 + f_2^2 - pf_1 + pf_2 = 0 \tag{39}$$

that defines a circle along which the performance is constant. To obtain the circle parameters in terms of performance, we write

$$f_1^2 + f_2^2 - pf_1 + pf_2 \equiv (x - x_1)^2 + (y - y_1)^2 - R^2 \tag{40}$$

From (40) we obtain the center coordinates x_1 ,y_1 and the radius R of the circle in terms of performance as

$$\begin{aligned} x_1 &= p/2 \\ y_1 &= p/2 \\ R &= p/\sqrt{2} \end{aligned} \tag{41}$$

The performance circle with the presence of three progressive Pareto fronts are schematically shown in figure 12. From this figure, it is seen that the maximum performance

is at the locations where the either objective is maximal at the Pareto front. If both objectives are equal, the maximal performance takes its lowest value and the degree of departing from the equality means a better performance in Pareto sense. This result is very significant since it reveals that, a design can have a better performance if some measured extremity in one way or other is exercised. It is meant that, if a better performance is obtained, then most presumably extremity will be observed in this design.

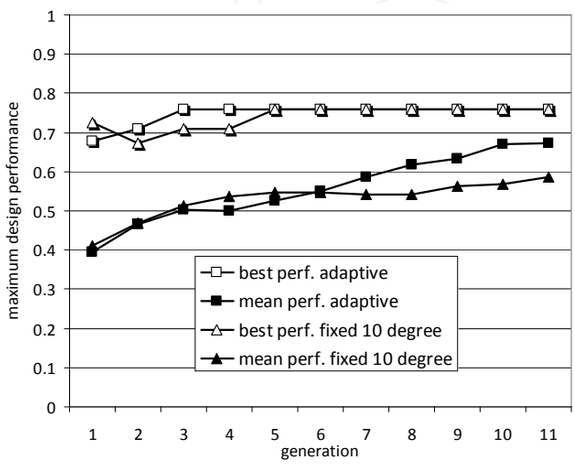


Fig. 13. Maximum design performance obtained during the multi-objective evolutionary search

3.3 Application results

The results from the design with multi-objective optimization are presented in figures 13, and 14-17. Figure 13 shows the maximum design performance p obtained using (39). From figure 13 it is seen that the average maximal performance is smoothly converging to its final value, whereas in the case a fixed angle is used, the performance is not increasing beyond a certain level after a few generations. The former result is a manifestation of the stochastic adaptive process. It indicates the superiority of the adaptive relaxation approach proposed with respect to non-adaptive relaxation.

To exemplify the solutions on the Pareto front, four resulting Pareto-optimal designs D1-D4 are shown in figures 14-17, respectively. The left part of the figure shows the instantiation of the solutions in decision space. The right part of the figure shows the same solution in objective space together with the other Pareto optimal solutions obtained. It is noted that in objective space a solution is represented by a sphere. The size of the sphere indicates the maximal performance value of the corresponding solution. That is, a large sphere indicates a high maximal design performance, and conversely a small sphere indicates a low performance.

Design D1 is the design among the Pareto solutions having the highest maximal design performance, as obtained by (37), namely p=.77. It has a high *form* and functional performance, namely .80, and .91 respectively, while its energy performance is moderate, being only .42. The high performance as to *form* is partially due to the elongated shape of the building body. The low energy performance is mainly due to the large surface of the building, and that offices O4 are exposed to the west side (lower part of the figure). These properties of the solution violate the energy requirements modelled in the fuzzy neural tree in figure 10.

Design D2 has a high energy performance (.83), while form and functionality are moderate (.54 and .41). Its maximal design performance is *p*=.65. The high energy performance is due to the more compact shape, and the fact that offices are not exposed in western direction, as required.

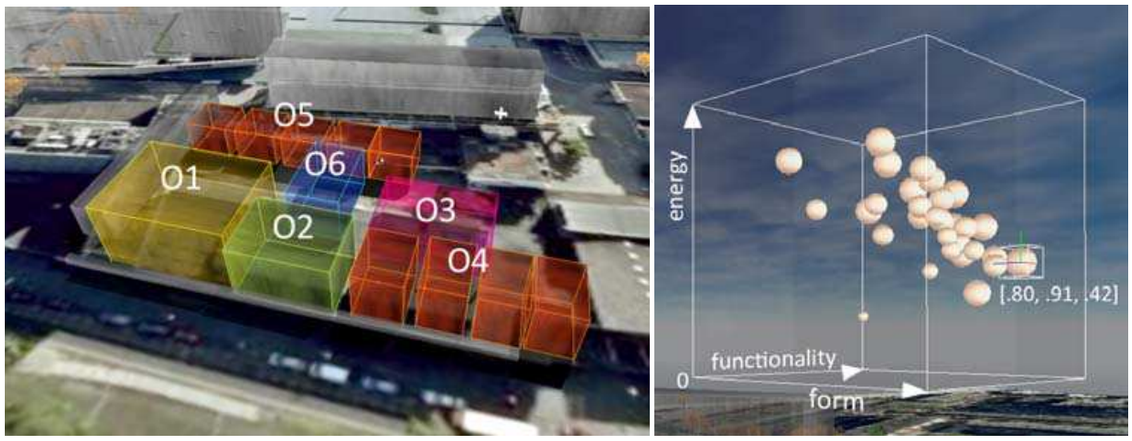


Fig. 14. Design D1 having maximal design performance

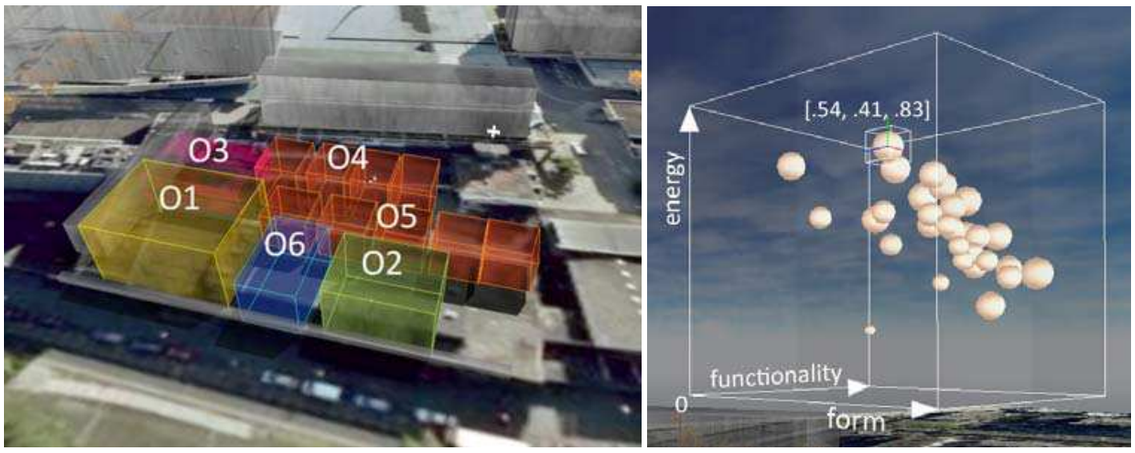


Fig. 15. Design D2 having a high energy performance

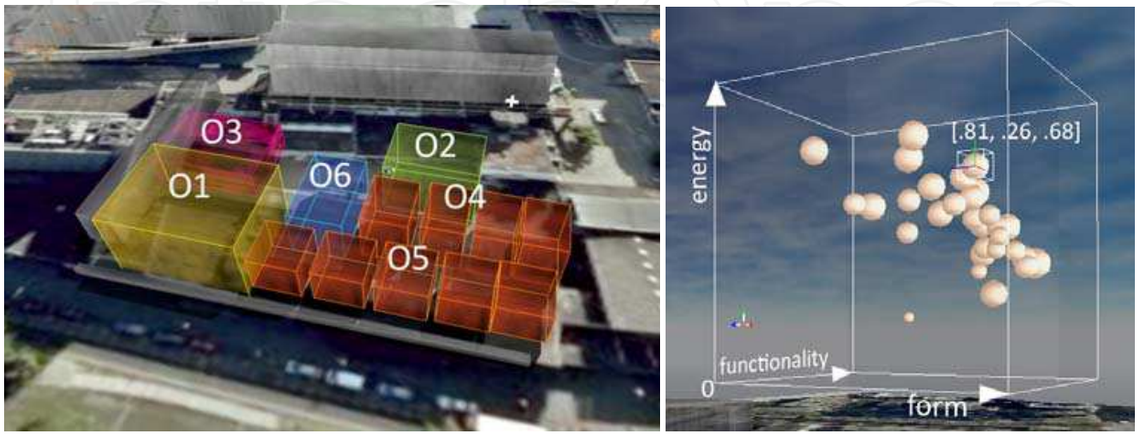


Fig. 16. Design D3 having a high form performance

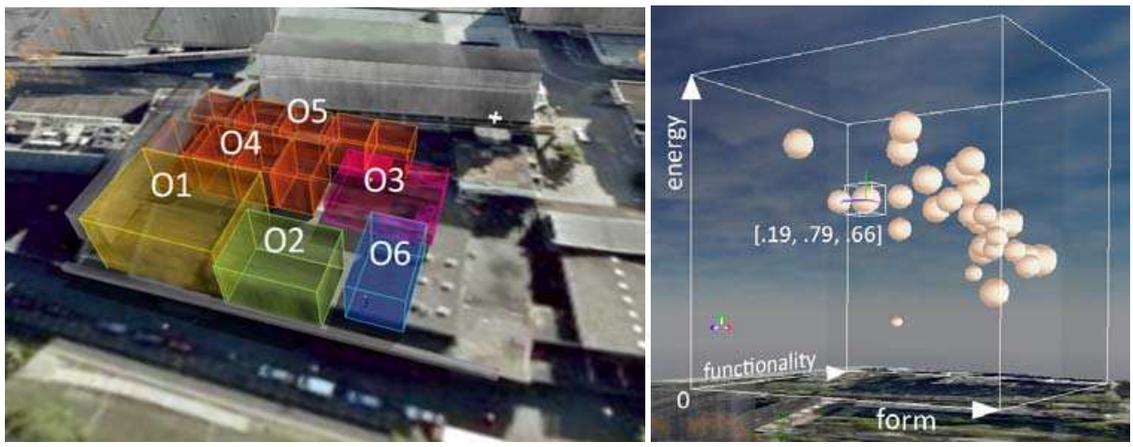


Fig. 17. Design D4 having a high functionality performance

Design D3 has a high form related performance (.81), while functionality is low (.26) and energy performance moderate (.68). Its maximal design performance is p=.67. The form related performance is high, because the building body is rather long and its façade to the west side consists of an office and the hotel building, which is fulfilling the form related preferences in this task. Functionality is low, for instance because the lobby O2 is located far from the street in the west, which is undesirable according to the requirements.

Design D4 has a high *functionality* (.79), while the form related performance is low (.19) and energy performance is moderate (.66). It design performance is p=.67. The high functionality is due to proximity among several objects, while hotel O1 and conference O3 are still quite distant from each other, as required. The low performance as to form is due to the square-like shape of the building.

From the results we note that all four Pareto solutions investigated, although being located at different extremities of the front, have shared properties. Namely the hotel object O1 is always located at the street corner. This means that this feature is desirable for any optimal solution, revealing a principle applicable to this design task. This discovery is referred to as *innovization* principle in the literature (Deb & Srinivasan, 2006).

From the results we note that design D1 has a maximal performance that is higher than the other Pareto optimal designs described by factor 1.15. That is, D1 clearly outperforms the others regarding their respective maximal performance. This means that when there is no a-priori bias for any of the three objectives, it is more proficient to be less concerned with energy, but to aim for maximal functionality and form qualities instead in the particular design task at hand. That is, in absence of second-order preferences, design D1 should be built, rather than the other designs.

4. Conclusions

A novel adaptive approach for formation of the Pareto front in multi-objective optimization is presented. The approach is an adaptive stochastic search, where a relaxed dominance concept is introduced, and the relaxation angle is adapted during the search. This approach is a dual counterpart of gradient-based stochastic adaptive algorithms. In this duality the fitnessfunction is the dual of stochastic gradient, and the degree of relaxation is the dual of the step-size parameter. The adaptation is found to be significantly favorable for

convergence and diversity in the multi-objective optimization. This is demonstrated in an application from architectural design, where a building consisting of several volumes is obtained. In this design the location of the volumes relative to each other is subject to optimization. This is accomplished by identifying an optimal sequence of arranging the volumes, so that three objectives pertaining to the building are satisfied. These objectives are functionality, energy aspects, and form related aspects. The linguistic nature of these requirements is treated by using a fuzzy neural tree approach, that is able to handle the imprecision and complexity inherent to the concepts, forming a model. This model plays the role of fitness function in the adaptive MOEA, so that the search is endowed with some human like reasoning during the search. The designs obtained are diverse, and have a high design performance. This performance is the maximum suitability one can attribute to a solution without preference for any objectives. The location where this maximal performance is constant in objective space constitutes a circle for two objectives, and a sphere for three or more objectives etc. Using this location information and associating it with the Pareto front, it is shown that a solution satisfying each objective approximately equally yields a minimal performance among the Pareto solutions. A Pareto solution deviating from this equality is preferred. This situation in two dimensional cases can metaphorically be referred to the preference for the *golden section* concept in design, in comparison to square shape. Namely a square is considered to be visually less appealing, whereas the golden section is more interesting having some directionality in it. In general this means that designs can have a better performance if some measured extremity in one way or other is exercised, provided there is no preference. The results from the application show that in the particular task at hand one should look for designs with outstanding functionality, rather than a design with outstanding energy performance. This is due to the shape of the Pareto front. The results have shown that the goal of adaptive formation of the Pareto front is satisfactorily achieved, and this is demonstrated.

Appendix - Coordinate Transform

Let us consider the elements of an n-dimensional numerical vector $\boldsymbol{x}=[x_1,x_2, \ldots ,x_n]$ whose components represent the directions of the n basic mutually orthogonal unit vectors

$$i_1 = [1,0,0..,0], \ldots\ldots \ , i_n = [0,0, \ldots . ,1] \tag{I}$$

which lie along the axes of reference rectangular coordinates x_1, x_2, …., x_n in n-dimensional space. The numerical vector $\boldsymbol{x}$ can be expressed by

$$x = [x_1, x_2, \ldots . , x_n] = \sum_{k=1}^{n} x_k i_k \tag{II}$$

Let a new coordinate system in n-dimensional space is choisen whose geometrical unit vectors i_1', i_2', … ,i_n' in the directions of axes of the new coordinates x_1', x_2', …., x_n' are linearly independent and are related to the original unit vectors by the equations

$$\begin{aligned} i_1^{'} &= q_{11}i_1 + q_{21}i_2 + \ldots , q_{n1}i_n \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ i_n^{'} &= q_{1n}i_1 + q_{2n}i_2 + \ldots , q_{nn}i_n \end{aligned} \tag{III}$$

which can be expressed as a matrix equation

$$\begin{bmatrix} i_1^{'} \\ \dots \\ i_n^{'} \end{bmatrix} = \begin{bmatrix} q_{11}\ q_{21}\ \dots\dots\ q_{n1} \\ \dots\dots\dots\dots \\ q_{1n}\ c_{2n}\ \dots\dots\ q_{nn} \end{bmatrix} \begin{bmatrix} i_1 \\ \dots\dots \\ i_n \end{bmatrix} \tag{IV}$$

From (III), one can write the set of equation in the following form.

$$i_k^{'} = \sum_{p=1}^{n} q_{pk} i_r \tag{V}$$

The geometrical vector determined by the components of the numerical vector x can be expressed by the components of the numerical vector x′ as follows

$$x^{'} = x_1^{'} i_1^{'} + x_2^{'} i_2^{'} + \dots , x_n^{'} = \sum_{k=1}^{n} x_k^{'} i_k^{'} \tag{VI}$$

Substitution of (IV) into (V) yields

$$x = \sum_{k=1}^{n} \sum_{p=1}^{n} x_k^{'} q_{pk} i_p = \sum_{p=1}^{n} \left(\sum_{k=1}^{n} q_{pk} x_k^{'} \right) i_p \tag{VII}$$

Comparison of (II) and (VII) yields

$$x_p = \sum_{k=1}^{n} q_{rk} x_k^{'} \tag{VIII}$$

The matrix equation form of (VIII) is given by

$$x = Q\, x^{'} \tag{IX}$$

where Q is the transformation matrix which is

$$Q = \begin{bmatrix} q_{11}\ q_{12}\ \dots\dots\ q_{1n} \\ \dots\dots\dots\dots \\ q_{n1}\ q_{n2}\ \dots\dots\ q_{nn} \end{bmatrix} \tag{X}$$

One should note that, Q is the transpose of the coefficient matrix of the equations in (IV). In order that the new unit vectors be linearly independent, and hence span n-dimensional space, the matrix Q must be non-singular. In this case from (IX) we can write

$$x^{'} = Q^{-1}\, x \tag{XI}$$

It may be interesting to note that, if the new unit vectors are mutually orthogonal, we obtain

$$Q^T Q = I \qquad or \qquad Q^T = Q^{-1} \tag{XII}$$

This result implies that Q is an *orthogonal matrix*.

5. References

Bazaraa, M. S.; Sherali, H. D. & Shetty, C. M. (1993), *Nonlinear Programming*, John Wiley & Sons, Inc, ISBN. 0-471-55793-5, New York

Branke, J.; Kaussler, T. & Schmeck, H. (2000), Guiding multi-objective evolutionary algorithms towards interesting regions, *Technical Report No: 399*, Institute AIFB University of Karlsruhe, Germany

Branke, J.; Kaussler, T. & Schmeck, H. (2001). Guidance in evolutionary multi-objective optimization, *Advances in Engineering Software*, vol. 32, pp. 499-507

Ciftcioglu, Ö. & Bittermann, M. S. (2008). Solution diversity in multi-objective optimization: A study in virtual reality, *Proc. World Congress on Computational Intelligence WCCI 2008*, Hong Kong, 2008,

Coello, C. A. C. (1999). An updated survey of GA-based multi-objective optimization techniques, *ACM Computing Surveys*, vol. 32, no. 2, pp. 109-143

Coello, C. A. C., Veldhuizen, D. A. & Lamont, G. B. (2003). *Evolutionary Algorithms for Solving Multiobjective Problems*, Kluwer Academic Publishers, Boston

Deb, K., Pratap, A., Agarwal, S. & Meyarivan, T. (2000). A fast and elitist multi-objective genetic algorithm: NSGA-II, *IEEE Transactions on Evolutionary Computation*, vol. 6, pp. 182-197

Deb, K. (2001). *Multiobjective Optimization Using Evolutionary Algorithms*, John Wiley & Sons

Deb, K., Zope, P. & Jain, A. (2003). Distributed computing of Pareto-optimal solutions with evolutionary algorithms, *Proc. 9th annual conference on Genetic and evolutionary computation,* pp. 532-549, 978-1-59593-697-4 London, England, ACM New York

Deb, K. & Srinivasan, A. (2006). Innovization: Innovating design principles through optimization, *Proc. of GECCO '06*, pp. 1629-1636, Seattle, Washington, July 8-12, 2006

Deb, K., Sundar, J., Bhaskara, U. & Chaudhuri, S. (2006). Reference point based multi-objective optimization using evolutionary algorithm, *Int. J. Comp. Intelligence Research*, vol. 2, no. 3, pp. 273-286

Eiben, A. E., Hinterding, R. & Michaelwicz, Z. (1999). Parameter control in evolutionary algorithms, *IEEE Trans. Evolutionary Computation*, vol. 3, no. 2, pp. 124-141

Farhang-Boroujeny, B. (1998), *Adaptive Filters*, John Wiley & Sons, New York

Horn, J., Nafploitis, N. & Goldberg, D. E. (1994). A niched pareto genetic algorithm for multiobjective optimization, *Proc. First IEEE Conf. on Evolutionary Computation*, pp. 82-87, 1994, IEEE Press

Hughes, E. J. (2005). Evolutionary many-objective optimisation: Many once or one many? *Proc. IEEE Congress on Evolutionary Computation CEC'2005*, pp. 222-227, Edinburgh, Scotland, 2005, IEEE Service Center

Jaszkiewicz, A. (2004). On the computational efficiency of multiple objective metaheuristics: The knapsack problem case study, *European Journal of Operational Research*, vol. 158, no. 2, pp. 418-433

Kuester, J. (1973). *Optimization Techniques with FORTRAN*, McGraw-Hill College, 0070356068

Laumanns, M., Thiele, L., Deb, K. & Zitzler, E. (2002). Combining convergence and diversity in evolutionary multiobjective optimization, *Evolutionary Computation*, vol. 10, no. 3, pp. 262-282

Mitra, S., Pal, S. K. & Mitra, P. (2002). Data mining in soft computing framework: A survey, *IEEE Trans. on Neural Networks*, vol. 13, no. 1, pp. 3-14

Sato, H., H.E. Aguirre, and K. Tanaka (2007). Controlling dominance area of solutions and its impact on the performance of MOEAs, *Lecture Notes in Computer Science, Evolutionary Multi-Criterion Optimization*, ISSN 0302-9743, Springer Berlin/Heidelberg, pp. 5-20, May

Yao, X., Liu, Y. & Lin, G. M. (1999). Evolutionary programming made faster, *IEEE Trans. Evolutionary Computation*, vol. 3, pp. 82-102

Zitzler, E. & Thiele, L. (1999). Multiobjective evolutionary algorithms: A comparative case study and the strength Pareto approach, *IEEE Trans. Evolutionary Computation*, vol. 3, no. 4, pp. 257-271

Evolutionary Computation

Edited by Wellington Pinheiro dos Santos

ISBN 978-953-307-008-7
Hard cover, 572 pages
Publisher InTech
Published online 01, October, 2009
Published in print edition October, 2009

This book presents several recent advances on Evolutionary Computation, specially evolution-based optimization methods and hybrid algorithms for several applications, from optimization and learning to pattern recognition and bioinformatics. This book also presents new algorithms based on several analogies and metafores, where one of them is based on philosophy, specifically on the philosophy of praxis and dialectics. In this book it is also presented interesting applications on bioinformatics, specially the use of particle swarms to discover gene expression patterns in DNA microarrays. Therefore, this book features representative work on the field of evolutionary computation and applied sciences. The intended audience is graduate, undergraduate, researchers, and anyone who wishes to become familiar with the latest research work on this field.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Ozer Ciftcioglu and Michael S. Bittermann (2009). Adaptive Formation of Pareto Front in Evolutionary Multi-objective Optimization, Evolutionary Computation, Wellington Pinheiro dos Santos (Ed.), ISBN: 978-953-307-008-7, InTech, Available from: http://www.intechopen.com/books/evolutionary-computation/adaptive-formation-of-pareto-front-in-evolutionary-multi-objective-optimization

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

IntechOpen