An Overview of Evolutionary Algorithms in Multiobjective Optimization

Carlos M. Fonseca* and Peter J. Fleming†
Dept. Automatic Control and Systems Eng.
University of Sheffield
Sheffield S1 4DU, U.K.
Tel.: +44 742 825236
Fax: +44 742 731729

July 15, 1994

Abstract

The application of evolutionary algorithms in multiobjective optimization, proposed relatively early in their history, is currently receiving growing interest from researchers with various backgrounds. The paper reviews and discusses current multiobjective evolutionary approaches, identifying some of the issues raised by how they handle multiple objectives, such as how they affect the fitness landscape. Directions for future research are identified from the discussion.

Keywords: evolutionary algorithms, multiobjective optimization, fitness assignment, search strategies, bibliography.

*E-mail: C.Fonseca@shef.ac.uk
†E-mail: P.Fleming@shef.ac.uk

1 Introduction

The simultaneous optimization of competing objective functions deviates from single function optimization in that it seldom admits a single, perfect solution. Instead, multiobjective optimization (MO) problems tend to be characterized by a family of alternatives which must be considered equivalent in the absence of information concerning the relevance of each objective relative to the others. Multiple solutions, or multimodality, arise in even the simplest non-trivial case of two competing unimodal convex objectives. As the number of competing objectives increases and less well-behaved objectives are considered, the problem rapidly becomes increasingly complex.

Early in their development, it was recognized that evolutionary algorithms (EAs) were possibly well-suited to multiobjective optimization. Multiple individuals can search for multiple solutions in parallel, eventually taking advantage of any similarities available in the solution set. The ability to handle complex problems, involving features such as discontinuities, multimodality, disjoint feasible spaces and noisy function evaluations, reinforces the potential effectiveness of EAs in multiobjective search and optimization.

The paper reviews current evolutionary approaches to multiobjective optimization, discussing their similarities and differences. It also tries to identify the main issues raised by multiobjective optimization in the context of evolutionary search, and how the methods discussed address them. From the discussion, directions for future work in multiobjective evolutionary algorithms will be identified.

2 Evolutionary approaches to multiobjective optimization

The solution set of a multiobjective optimization problem consists of all those vectors such that their components cannot be all simultaneously improved. This is known as the concept of Pareto optimality, and the solution set is known as the Pareto-optimal set. Pareto-optimal solutions are also called non-dominated, or non-inferior, solutions. In practice, however, a single compromise solution is usually sought; it is selected from the Pareto-optimal set according to some preference information.

Because evolutionary algorithms require scalar fitness information to work on, a scalarization of the objective vectors is necessary. In most problems where no global criterion directly emerges from the problem formulation, objectives are often artificially combined, or aggregated, into a scalar function according to some understanding of the problem, and the EA applied. Many such approaches developed for use with conventional optimizers can also be used with EAs.

Optimizing a combination of the objectives has the advantage of producing a single compromise solution, requiring no further interaction with the decision maker (DM). The problem is, if the optimal solution cannot be accepted, either due to the function used excluding aspects of the problem which were unknown prior to optimization or to an inappropriate setting of the coefficients of the combining function, new runs of the optimizer may be required until a suitable solution is found.

Several applications of evolutionary algorithms in the optimization of aggregating functions have been reported in the literature, from the simple weighted sum approach (Jakob *et al.*, 1992) to target vector optimization (Wienke *et al.*, 1992). Goal attainment, among other methods, was used by Wilson and Macleod (1993), who also monitored the population for non-dominated solutions.

Handling constraints with penalty functions is yet another example of an additive aggregating function. The fact that penalty functions are generally problem dependent and, as a consequence, difficult to set has prompted the development of alternative approaches based on ranking (Powell and Skolnick, 1993).

2.1 Non-Pareto approaches

The first step towards treating objectives separately in EAs was given by Schaffer (1985). In his approach, known as the Vector Evaluated Genetic Algorithm (VEGA), appropriate fractions of the next generation, or sub-populations, were selected from the whole of the old generation according to each of the objectives, separately. Crossover and mutation were applied as usual after shuffling all the sub-populations together. Non-dominated individuals were identified by monitoring the population as it evolved.

Shuffling and merging all sub-populations corresponds, however, to averaging the fitness components associated with each of the objectives. In fact, the expected total number of offspring produced by each parent becomes the sum of the expected numbers of offspring produced by that parent according to each objective. Since Schaffer used proportional fitness assignment, these were in turn, proportional to the objectives themselves. The resulting expected fitness corresponded, therefore, to a linear function of the objectives where the weights depended on the distribution of the population at each generation (Richardson *et al.*, 1989; Schaffer, 1993). As a consequence, different non-dominated individuals were generally assigned different fitness values.

In the case of concave trade-off surfaces, the population tended to split into species particularly strong in each of the objectives. This can be understood by noting that points in concave regions of the trade-off cannot be found by optimizing a linear combination of the objectives, for any set of weights.

Fourman (1985) also addressed multiple objectives in a non-aggregating manner. Selection was performed by comparing pairs of individuals, each pair according to one of the objectives. In a first version of the algorithm, objectives were assigned different priorities by the user and individuals compared according to the objective with the highest priority. If this resulted in a tie, the objective with the second highest priority was used, and so on. This is known as the *lexicographic* ordering (Ben-Tal, 1980).

A second version, reported to work surprisingly well, consisted of randomly selecting the objective to be used in each comparison. Similarly to VEGA, this corresponds to averaging fitness across fitness components, each component being weighted by the probability of each objective being chosen to decide each tournament. However, the use of pairwise comparisons makes it essentially *different* from a linear combination of the objectives, because scale information is ignored. Thus, the population may still see as convex a trade-off surface actually concave, depending on its current distribution and, of course, on the problem.

Kursawe (1991) formulated a multiobjective version of evolutionary strategies (ESs). Once again, selection consisted of as many steps as there were objectives. At each step, one objective was selected randomly according to a probability vector, and used to dictate the deletion of an appropriate fraction of the current population. After selection, μ survivors became the parents of the next generation. A picture of the trade-off surface was produced from the points evaluated during the run. The non-stationary environment imposed by the selection procedure made the use of diploid individuals necessary.

Finally, and still based on the weighted sum approach, Hajela and Lin (1992) exploited the explicit parallelism provided by a population-based search by explicitly including the weights in the chromosome and promoting their diversity in the population through fitness sharing. As a consequence, one family of individuals evolved for each weight combination, concurrently.

2.2 Pareto-based approaches

Pareto-based fitness assignment was first proposed by Goldberg (1989), the idea being to assign equal probability of reproduction to all non-dominated individuals in the population. The method consisted of assigning rank 1 to the non-dominated individuals and removing them from contention, then finding a new set of non-dominated individuals, ranked 2, and so forth.

Fonseca and Fleming (1993) have proposed a slightly different scheme, whereby an individual's rank corresponds to the number of individuals in the current population by which it is dominated. Non-dominated individuals are, therefore, all assigned the same rank, while dominated ones are penalized according to the population density of the corresponding region of the trade-off surface.

By combining Pareto dominance with partial preference information in the form of a goal vector, they have also provided a means of evolving only a given region of the trade-off surface. While the basic ranking scheme remains unaltered, the Pareto comparison of the individuals selectively excludes those objectives which already satisfy their goals. Specifying fully unattainable goals causes objectives never to be excluded from comparison, which corresponds to the original Pareto ranking. Changing the goal values during the search alters the fitness landscape accordingly and allows the decision maker to direct the population to zoom in on a particular region of the trade-off surface.

Tournament selection based on Pareto dominance has also been proposed by Horn and Nafpliotis (1993). In addition to the two individuals competing in each tournament, a number of other individuals in the population was used to help determine whether the competitors were dominated or not. In the case where both competitors were either dominated or non-dominated, the result of the tournament was decided through sharing (see below). Population sizes considerably larger than usual were used so that the noise of the selection method could be tolerated by the emerging niches in the population.

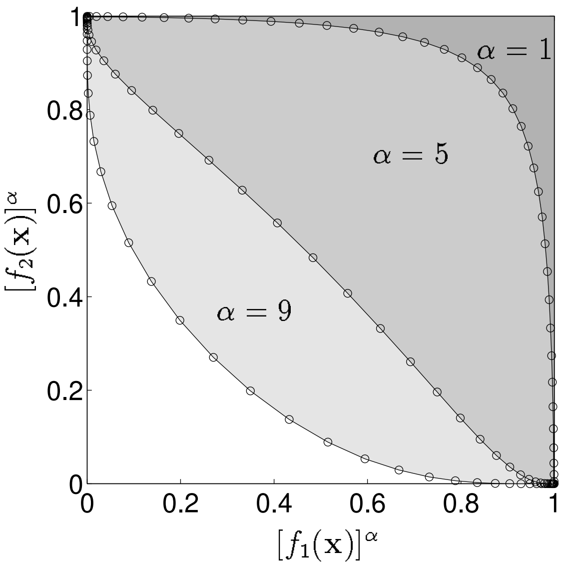

Figure 1: The concavity of the trade-off set is related to how the objectives are scaled.

The convexity of the trade-off surface depends on how the objectives are scaled. Non-linearly rescaling the objective values may convert a concave surface into a convex one, and vice-versa, as illustrated in Figure 1. The darker surface is the original, concave trade-off surface, corresponding to plotting $f_1(\mathbf{x})$ against $f_2(\mathbf{x})$, where $\mathbf{x}$ denotes the vector of free variables. The lighter surfaces correspond to plotting $[f_1(\mathbf{x})]^\alpha$ against $[f_2(\mathbf{x})]^\alpha$, for $\alpha = 5$ and $\alpha = 9$, the latter being clearly convex. Nevertheless, all are formulations of the same minimization problem which admit exactly the same solution set in phenotypic space.

Since order is preserved by monotonic transformations such as these, Pareto-ranking is blind to the convexity or the non-convexity of the trade-off surface. This is not to say that Pareto-ranking always precludes speciation. Speciation can still occur if certain regions of the trade-off are simply easier to find than others, but Pareto-ranking does eliminate sensitivity to the possible non-convexity of the trade-off.

A second possible advantage of Pareto-ranking is that, because it rewards good performance in any objective dimension regardless of the others, solutions which exhibit good performance in many, if not all, objective dimensions are more likely to be produced by recombination. This argument, which assumes some degree of independence between objectives, was already hinted at by Schaffer in his VEGA work and has been noted in more detail by Louis and Rawlins (1993). While Pareto-based selection may help find utopian solutions if they exist, multiobjective optimization would be unnecessary in those circumstances. In any case, the argument may still hold in the initial stages of the search.

2.3 Niche induction techniques

Pareto-based ranking correctly assigns all non-dominated individuals the same fitness, but that, on its own, does not guarantee that the Pareto set be uniformly sampled. When presented with multiple equivalent optima, finite populations tend to converge to only one of them, due to stochastic errors in the selection process. This phenomenon, known as genetic drift, has been observed in natural as well as in artificial evolution, and can also occur in Pareto-based evolutionary optimization.

The additional use of fitness sharing (Goldberg and Richardson, 1987) was proposed by Goldberg to prevent genetic drift and to promote the sampling of the whole Pareto set by the population. Fonseca and Fleming (1993) implemented fitness sharing in the objective domain and provided theory for estimating the necessary niche sizes based on the properties of the Pareto set. Horn and Nafpliotis (1993) also arrived at a form of fitness sharing in the objective domain. In addition, they suggested the use of a metric combining both the objective and the decision variable domains, leading to what they called nested sharing.

The viability of mating is another aspect which becomes relevant as the population distributes itself around multiple regions of optimality. Different regions of the trade-off surface may generally have very different genetic representations, which, to ensure viability, requires mating to happen only locally (Goldberg, 1989). So far, mating restriction has been implemented based on the distance between individuals in the objective domain, either directly (Fonseca and Fleming, 1993) or indirectly (Hajela and Lin, 1992).

3 Discussion

The handling of multiple objectives strongly interacts with evolutionary computation on many fronts, raising issues which can generally be accommodated in one of two broad classes, fitness assignment and search strategies.

3.1 Fitness assignment

The extension of evolutionary algorithms to the multiple objective case has mainly been concerned with multiobjective fitness assignment. According to how much preference information is incorporated in the fitness function, approaches range from complete preference information given, as when combining objective functions directly or prioritizing them, to no preference information given, as with Pareto-based ranking, and include the case where partial information is provided in order to restrict the search to only part of the Pareto set. Progressive refinement of partial preferences is also possible with EAs.

Independently of how much preference information is provided, the assigned fitness reflects a decision maker's understanding of the quality, or *utility*, of the points under assessment. Each selection step of an EA can be seen as a decision making problem involving as many alternatives as there are individuals in the population.

The effect of different fitness assignment strategies on the fitness landscape can be more easily understood by means of an example. Consider the simple bi-objective problem of simultaneously minimizing

$$\begin{aligned} f_1(x_1, x_2) &= 1 - \exp\left(-(x_1 - 1)^2 - (x_2 + 1)^2\right) \\ f_2(x_1, x_2) &= 1 - \exp\left(-(x_1 + 1)^2 - (x_2 - 1)^2\right) \end{aligned}$$

If individuals are ranked according to how many members of the population outperform them (Fonseca and Fleming, 1993), the ranking of a large, uniformly distributed population, normalized by the population size, can be interpreted as an estimate of the fraction of the search space which outperforms each particular point considered (global optima should be ranked zero). This applies equally to single-objective ranking.

Plotting the normalized ranks against the decision variables, x_1 and x_2 in this case, produces an anti-fitness, or cost, landscape, from which the actual

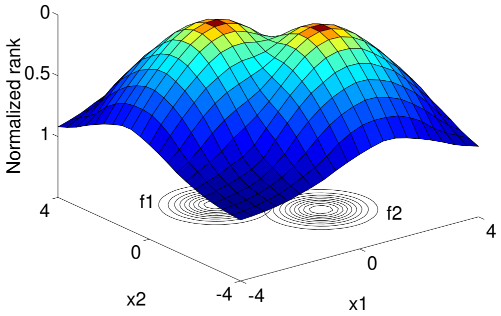

Figure 2: The cost landscape defined by ranking the sum of the objectives (The contour plots are those of the individual objective functions f_1 and f_2)

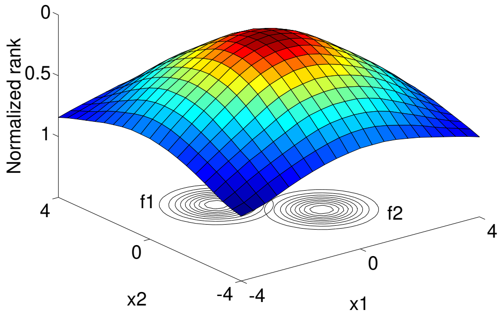

Figure 3: The cost landscape defined by ranking objectives separately and averaging the ranks

fitness landscape can be inferred. Cost landscapes for the example above are shown in Figures 2 to 5, corresponding to four different fitness assignment strategies based on ranking.

Figure 2 illustrates the single-objective ranking of the sum of the two objectives. The two peaks arise from the problem exhibiting a concave trade-off surface, and create the scope for speciation, or even genetic drift, to occur during evolutionary search.

In Figure 3, the average of the ranks computed according to each of the two objectives is shown. In this case, a single peak is located towards the middle of the Pareto-optimal set, and the concavity of the trade-off surface is no longer apparent. Binary tournaments according to one objective drawn at random can be expected to define a similar landscape.

Figure 4 shows the cost landscape for the ranking of the maximum of the two objectives, a simple case of goal programming. The single-peak is located on a non-smooth ridge, which makes gradient-based optimization difficult, if not impossible. For this reason, the goal attainment method as proposed by Gembicki (1974) is usually preferred to this approach.

Finally, in Figure 5, Pareto-ranking is used. Note how the Pareto-optimal set defines a flat ridge in the cost landscape.

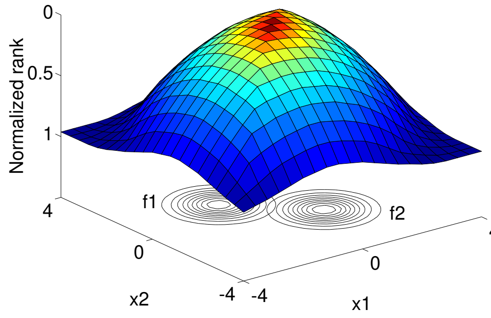

Figure 4: The cost landscape defined by ranking the maximum of the two objectives

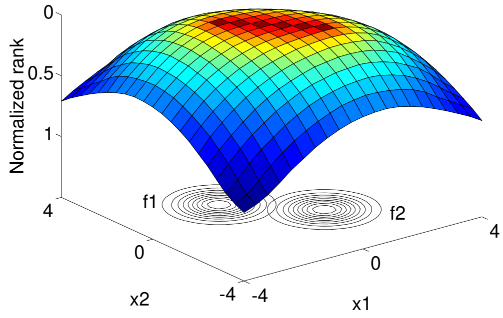

Figure 5: The cost landscape defined by Pareto-ranking

3.2 Search strategies

The ridges defined in the fitness landscape by Pareto-ranking and/or minimax approaches may not be parallel to any of the decision variable axes, or even straight! Although ridges, or equivalently, valleys, need not occur in single-objective optimization (Mühlenbein and Schlierkamp-Voosen, 1993), they do appear in this context, and can certainly be expected in almost any multiobjective problem.

Flat ridges raise two problems already encountered with other types of multimodality. Firstly, genetic drift may lead to the poor sampling of the solution set. This problem can be reasonably addressed by performing fitness-sharing based on a good closeness measure. Secondly, mating of well-performing individuals far apart from one another tends not to be viable. Mating restriction, or the absence of mating altogether, interprets the individuals populating the Pareto-front as a continuum of species. It seeks to reduce the formation of lethals by encouraging the formation of offspring similar to their parents, which means a less exploratory search. This has been the approach used so far in Pareto-based search.

The alternative interpretation of the Pareto-set as a genetically similar and, therefore, reproductively viable family of points would require the search for a suitable genetic representation in addition to the solution itself, because the location of the optima is not known prior to optimization. A fixed genetic representation also produces a reproductively viable family of points, but it does not necessarily correspond to the Pareto-set.

Ridges impose a second type of difficulty. Theoretical results by Wagner (1988) show that, under biologically reasonable assumptions, the rate of progression of unconstrained phenotypes on certain types of ridge-shaped landscapes is bounded, in which case it decreases rapidly as the number of decision variables increases. Fast progression cannot be achieved unless the genetic operators tend to produce individuals which stay inside the corridor. The self-adaptation of mutation variances and correlated mutations (1991), as implemented in evolution strategies, addresses this same problem, but has not yet been tried in Pareto-based search. Binary mutation, as usually implemented in genetic algorithms, can be particularly destructive if the ridge expresses a strong correlation between decision variables.

Finally, multiobjective fitness landscapes become non-stationary once the DM is allowed to interact with the search process and change the current preferences. Diploidy has already revealed its importance in handling non-stationary environments. Other relevant work is the combination of evolutionary and pure random search proposed by Grefenstette (1992).

4 Future perspectives

As discussed in the previous section, the EA can be seen as a sequence of decision making problems, each involving a finite number of alternatives. Current decision making theory, therefore, can certainly provide many answers on how to perform multiobjective selection in the context of EAs.

On the other hand, progress in decision making has always been strongly dependent on the power of the numerical techniques available to support it. Certain decision models, although simple to formulate, do not necessarily lead to numerically easy optimization problems (Dinkelbach, 1980). By easing the numerical difficulties inherent to other optimization methods, evolutionary algorithms open the way to the development of simpler, if not new, decision making approaches.

A very attractive aspect of the multiobjective evolutionary approach is the production of useful intermediate information which can be used by an intelligent DM to refine preferences and terminate the search upon satisfaction. In fact, the DM is not only asked to assess individual performance, but also to adjust the current preferences in the search for a compromise between the ideal and the possible in a limited amount of time. Goal setting, for example, is itself the object of study (Shi and Yu, 1989). This is an area where combinations of EAs and other learning paradigms may be particularly appropriate.

As far as the search strategy is concerned, much work has certainly yet to be done. In particular the emergence of niches in structured populations (Davidor, 1991) suggests the study of such models in the multiobjective case. The development of adaptive representations capable of capturing and exploiting directional trends in the fitness landscape, well advanced in the context of ESs, and/or the corresponding operators, is another important avenue for research. Combinations of genetic search and local optimization resulting in either Lamarckian or developmental Baldwin learning (Gruau and Whitley, 1993) may also provide a means of addressing the difficulties imposed by ridge-shaped landscapes.

Acknowledgement

The first author gratefully acknowledges support by Programa CIENCIA, Junta Nacional de Investigação Científica e Tecnológica, Portugal.

References

Bäck, T., Hoffmeister, F., and Schwefel, H.-P. (1991). A survey of evolution strategies. In (Belew and Booker, 1991), pages 2–9.

Belew, R. K. and Booker, L. B., editors (1991). *Genetic Algorithms: Proceedings of the Fourth International Conference.* Morgan Kaufmann, San Mateo, CA.

Ben-Tal, A. (1980). Characterization of Pareto and lexicographic optimal solutions. In (Fandel and Gal, 1980), pages 1–11.

Davidor, Y. (1991). A naturally occurring niche and species phenomenon: The model and first results. In (Belew and Booker, 1991), pages 257–263.

Dinkelbach, W. (1980). Multicriteria decision models with specified goal levels. In (Fandel and Gal, 1980), pages 52–59.

Fandel, G. and Gal, T., editors (1980). *Multiple Criteria Decision Making Theory and Application*, volume 177 of *Lecture Notes in Economics and Mathematical Systems*. Springer-Verlag, Berlin.

Fonseca, C. M. and Fleming, P. J. (1993). Genetic algorithms for multi-objective optimization: Formulation, discussion and generalization. In (Forrest, 1993), pages 416–423.

Forrest, S., editor (1993). *Genetic Algorithms: Proceedings of the Fifth International Conference*. Morgan Kaufmann, San Mateo, CA.

Fourman, M. P. (1985). Compaction of symbolic layout using genetic algorithms. In (Grefenstette, 1985), pages 141–153.

Gembicki, F. W. (1974). *Vector Optimization for Control with Performance and Parameter Sensitivity Indices*. PhD thesis, Case Western Reserve University, Cleveland, Ohio, USA.

Goldberg, D. E. (1989). *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley, Reading, Massachusetts.

Goldberg, D. E. and Richardson, J. (1987). Genetic algorithms with sharing for multimodal function optimization. In Grefenstette, J. J., editor, *Genetic Algorithms and Their Applications: Proceedings of the Second International Conference on Genetic Algorithms*, pages 41–49. Lawrence Erlbaum.

Grefenstette, J. J., editor (1985). *Genetic Algorithms and Their Applications: Proceedings of the First International Conference on Genetic Algorithms*. Lawrence Erlbaum.

Grefenstette, J. J. (1992). Genetic algorithms for changing environments. In (Männer and Manderick, 1992), pages 137–144.

Gruau, F. and Whitley, D. (1993). Adding learning to the cellular developmentof neural networks: Evolution and the baldwin effect. *Evolutionary Computation*, 1(3):213–233.

Hajela, P. and Lin, C.-Y. (1992). Genetic search strategies in multicriterion optimal design. *Structural Optimization*, 4:99–107.

Horn, J. and Nafpliotis, N. (1993). Multiobjective optimization using the niched Pareto genetic algorithm. IlliGAL Report 93005, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA.

Jakob, W., Gorges-Schleuter, M., and Blume, C. (1992). Application of genetic algorithms to task planning and learning. In (Männer and Manderick, 1992), pages 291–300.

Kursawe, F. (1991). A variant of evolution strategies for vector optimization. In Schwefel, H.-P. and Männer, R., editors, *Parallel Problem Solving from Nature, 1st Workshop, Proceedings*, volume 496 of *Lecture Notes in Computer Science*, pages 193–197. Springer-Verlag, Berlin.

Louis, S. J. and Rawlins, G. J. E. (1993). Pareto optimality, GA-easiness and deception. In (Forrest, 1993), pages 118–123.

Männer, R. and Manderick, B., editors (1992). *Parallel Problem Solving from Nature, 2*. North-Holland, Amsterdam.

Mühlenbein, H. and Schlierkamp-Voosen, D. (1993). Predictive models for the breeder genetic algorithm i. continuous parameter optimization. *Evolutionary Computation*, 1(1):25–49.

Powell, D. and Skolnick, M. M. (1993). Using genetic algorithms in engineering design optimization with non-linear constraints. In (Forrest, 1993), pages 424–431.

Richardson, J. T., Palmer, M. R., Liepins, G., and Hilliard, M. (1989). Some guidelines for genetic algorithms with penalty functions. In Schaffer, J. D., editor, *Proceedings of the Third International Conference on Genetic Algorithms*, pages 191–197. Morgan Kaufmann, San Mateo, CA.

Schaffer, J. D. (1985). Multiple objective optimization with vector evaluated genetic algorithms. In (Grefenstette, 1985), pages 93–100.

Schaffer, J. D. (1993). Personal communication.

Shi, Y. and Yu, P. L. (1989). Goal setting and compromise solutions. In Karpak, B. and Zionts, S., editors, *Multiple Criteria Decision Making* *and Risk Analysis Using Microcomputers*, volume 56 of *NATO ASI Series F: Computer and Systems Sciences*, pages 165–204. Springer-Verlag, Berlin.

Wagner, G. P. (1988). The influence of variation and of developmental constraints on the rate of multivariate phenotypic evolution. *Journal of Evolutionary Biology*, 1(1):45–66.

Wienke, D., Lucasius, C., and Kateman, G. (1992). Multicriteria target vector optimization of analytical procedures using a genetic algorithm. Part I. Theory, numerical simulations and application to atomic emission spectroscopy. *Analytica Chimica Acta*, 265(2):211–225.

Wilson, P. B. and Macleod, M. D. (1993). Low implementation cost IIR digital filter design using genetic algorithms. In *IEE/IEEE Workshop on Natural Algorithms in Signal Processing*, volume 1, pages 4/1–4/8, Chelmsford, U.K.