Evolution Strategies for Vector Optimization

Frank Kursawe

University of Dortmund
Department of Computer Science
Chair of Systems Analysis
44221 Dortmund, Germany
E-Mail: kursawe@LS11.informatik.uni-dortmund.de
Phone: +49-231-755-2622, Fax: +49-231-755-2047

Abstract

Evolution strategies — a stochastic optimization method originally designed for single criterion problems — have been modified in such a way that they can also tackle multiple criteria problems. Instead of computing only one efficient solution interactively, a decision maker can collect as many members of the Pareto set as needed before making up his mind.
Apart from this feature one could also reflect upon the algorithm as a simple model of biological evolution. Following this idea one might emphasize the algorithm's capability of self–adapting its parameters. Furthermore, the effect of polyploid individuals corresponds in both 'worlds'.

1 Introduction

It has become increasingly obvious that the optimization under a single scalar–valued criterion — often a monetary one — fails to reflect the variety of aspects in a world getting more and more complex. Although V. Pareto [4] laid the mathematical foundations already about a hundred years ago the existing tools for multiple criteria decision making often demand too much of non–mathematicians who want to use them.

In order to overcome these difficulties a new method based on evolution strategies has been developed being capable of giving a good insight into the structure of the Pareto set by computing a finite number of efficient solutions.

2 Shortcomings of Conventional Methods

Numerous methods have been developed for vector optimization, e.g. linear vector optimization, game theory, one global replacement criterion and others [3, 5]. Many of these algorithms reduce the problem formulation to a scalar one. This seems tempting for two reasons:

- In contrast to the original problem the reduced one (hopefully) has one distinct solution.
- The decision maker may choose from a variety of approved procedures from the domain of single criterion optimization.

Like many other approaches, however, this idea reduces the decision space prematurely, i.e. before enough information is available.

Despite of the large number of numerical tools for vector optimization problems several questions still remain:

- If a utility function is assumed, does it exist? And if so, has it been understood correctly? How does the chosen utility function influence the solution? How can correlated objectives be dealt with? Is the solution obtained efficient? Furthermore, does it make sense to obliterate the information obtainable from K objectives?
- If the decision maker has to deliver aspiration levels, how can one measure their influence on the solution? How can one deal with the fact that any metric discriminates those objectives with small absolute values? How does the solution obtained depend on the chosen metric?
- What is there to do if the objectives are not commensurable?

3 Evolution Strategies ...

3.1 ... for Single Criterion Optimization

Evolution can be regarded as a sequence of self–organization steps, i.e. as the underlying universal principle of any kind of self–organization. Modern research has proved nature's strategies worth copying for technical or numerical optimization. Pioneering work in this direction has been done by Holland [2], Bremermann [1], Rechenberg [6] and Schwefel [7].

Technical problems have led to the development of *evolution strategies* as a method for experimental optimization. Nowadays, they can also deal with parameter optimization problems given as mathematical models of the type

$$\boxed{\min\{f(\underline{x}) \mid \underline{x} \in M \subseteq I\!R^n\}}.$$

Multi–membered evolution strategies were first proposed by Schwefel [7, 8] as a robust, general purpose optimization algorithm being very modest in terms of prerequisite mathematical assumptions: One only needs a criterion determining whether one alternative is better than another. As well as genetic algorithms they have shown to be capable of searching for the global optimum in parameter spaces which cause difficulties for gradient algorithms. Their range of application also covers NP–hard problems such as the Travelling–Salesman–Problem or problems with an optimum changing over time. Even for pattern matching which does not belong to the classical domain of optimization evolution strategies turned out to be useful. In fact, one can treat all problems which provide some criterion (environment) evaluating an individual's fitness. Regarding the objective function as a 'black box' one can even search for optimum states in large–scale simulation models.

In order to solve these single criterion optimization problems it proved to be sufficient to imitate the following principles of nature:

- population (in order to enable collective learning)
- haploid individuals
- synchronous generations
- sexual propagation with recombination/crossing–over
- random mating
- mutation ('driving force')
- selection ('steering wheel')

Generally, one can understand mutation as a process of varying or generating propositions. Mutations in evolution strategies consist of adding normally distributed random numbers with expectation zero and standard deviation σ_i to the object (decision) variables (x_i) thus securing a certain similarity between a parent and its offspring(-s). Mutations of the strategy (stepsize) variables σ_i, however, are log–normally distributed so that doubling and halving may occur with equal probability. The selection step then evaluates the usefulness of these variations.

3.2 ... for Multiple Criteria Optimization

For multiple criteria problems, however, two of the principles mentioned above have to be modified:

- Since the environment now consists of K objectives the selection step is provided with a fixed user–definable vector that determines the probability of each objective to become the sorting criterion in the K iterations of the selection loop. Alternatively, this vector may be allowed to change randomly over time.
- Furthermore, the extension of an individual's genes by recessive information turned out to be necessary in order to maintain the population's capability of coping with a changing environment. The recessive genes enable a fast reaction after a sudden variation of the probability vector. One can also observe this behaviour in nature: The younger the environment the higher the portion of polyploid organisms. Figure 1 illustrates the interior structure of a diploid individual:

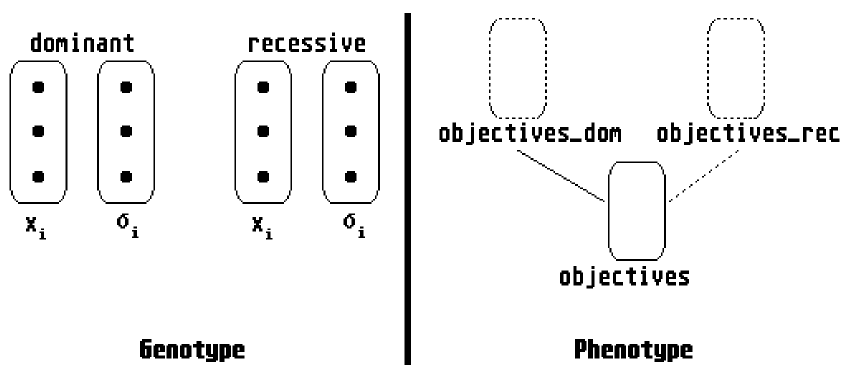

Figure 1: Genotype / Phenotype of an Individual

By presenting many solutions the algorithm provides the user with an idea of the tradeoffs between the objectives. It should be noted that efficient solutions in one generation may become dominated by individuals emerging in a later generation. This explaines the non–efficient points in figure 2 (left). For efficiency reasons the 'parents' of the next generation are stored provisionally in an array that is cleaned out if there is not enough space left for further individuals. If this operation does not result in enough free space solutions 'close' to one another are deleted. As an important side effect the elements of the Pareto set are forced apart thus allowing a good survey with only a finite number of solutions. Figure 2 (right) displays the situation after tidying up:

$$f_1(\underline{x}) = \sum_{i=1}^{n}(-10\, e^{-0.2\sqrt{x_i^2+x_{i+1}^2}}), \qquad f_2(\underline{x}) = \sum_{i=1}^{n}(|x_i|^{0.8} + 5\,\sin(x_i)^3)$$

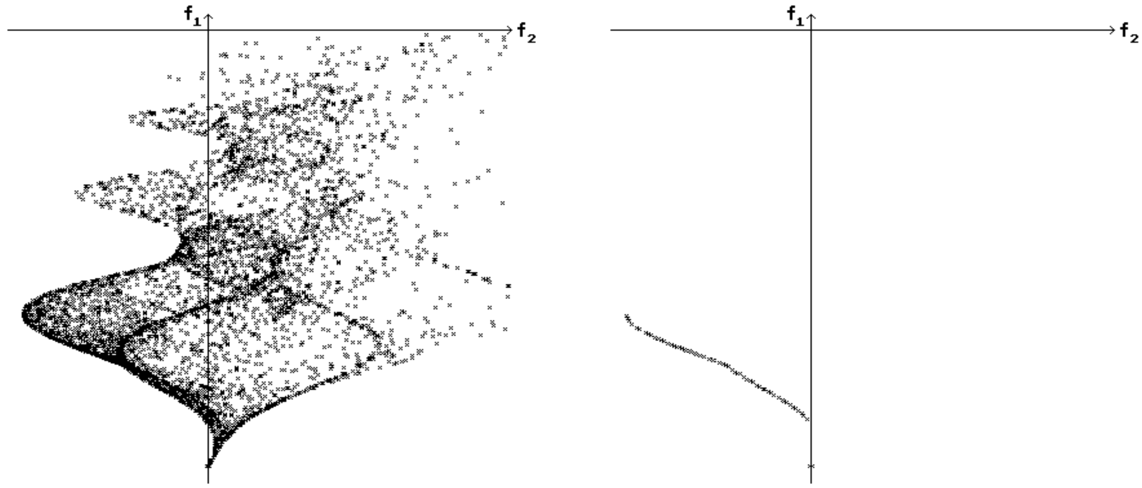

Figure 2: Graphical Output of the Algorithm

4 Results

When working with diploid individuals the inclusion of the recessive genes in the selection step turns out to be vital. Otherwise, undisturbed by the outside world they lead such a life of their own that an individual whose dominant genes have been freshened up with recessive material has no chance of surviving the next selection step. The best results were achieved with a probability of about 1/3 for exchanging dominant and recessive genes. This value also serves as a factor when putting together the overall fitness vector. Only in this way the additional recessive material can serve as a stock of variants. From further test runs one can also conclude that diploid or, in general, polyploid individuals are not worth the additional computing time in a static environment consisting only of one objective function.

Since the algorithm tries to cover the Pareto set as good as possible a probability distribution forcing certain minimum changes during the mutation step ought to yield better results. Indeed, the (symmetric) Weibull distribution turned out to be better than the Gaussian distribution.

The stochastic approach towards vector optimization problems via evolution strategies leads to one major advantage: In contrast to other methods no subjective decisions are required during the course of the iterations. Instead of narrowing the control variables space or the objective space by deciding about the future direction of the search from an 'information vacuum' [3] the decision maker can collect as much information as needed before making a choice which of the alternatives should be realized. Moreover, using a population while looking for a set of efficient solutions seems to be more appropriate than just trying to improve one 'current best' solution.

One might exploit the algorithm's capability of self–adapting its parameters even further: The exchange rate between dominant and recessive genetic material can be adjusted on–line thus providing the user with a measure of convergence. The self–adaptation property largely depends on a selection scheme that forces the algorithm to 'forget' the good solutions ('parents') of one generation. When accepting a possible recession from one generation to the next on the phenotype level individuals with a better 'model' of their environment, i.e. better step sizes σ_i are likely to emerge in later generations. This kind of selection seems to be lavish at first sight but it favours better adapted settings, thus speeding up the search in the long run.

5 Outlook

In future research it will be interesting to see whether further principles of nature are worth copying, such as aging, fertility rates depending on the relative fitness or parallel (sub–)populations. For example, the selection loop could be modified in the following way: Each time the appropriate fraction of the next generation is selected according to all elements of the fitness vector one after another. By doing so one guarantees the survival of the best individuals on each objective and, simultaneously, enhances the reproduction probability of those individuals selected more than once, i.e. those that are better than average on more than one objective.

One should, however, carry out 'mutations' of the algorithm carefully and only if the underlying natural principle has been fully understood.

References

[1] H. J. Bremermann, *Optimization through evolution and recombination*, in Self–Organizing Systems, M. C. Yovits, G. T. Jacobi, and D. G. Goldstein, eds., Spartan Books, Washington, D.C., 1962, pp. 93–106.

[2] J. H. Holland, *Adaptation in Natural and Artificial Systems*, University of Michigan Press, Ann Arbor, Michigan, 1975.

[3] C.-L. Hwang and A. S. M. Masud, *Multiple Objective Decision Making — Methods and Applications*, Springer, Berlin, 1979.

[4] V. Pareto, *Cours d'Économie Politique*, Lausanne, 1896.

[5] M. Peschel, *Modellbildung und Steuerung mit Hilfe der Polyoptimierung*, VEB Verlag Technik, Berlin, 1980.

[6] I. Rechenberg, *Evolutionsstrategie: Optimierung technischer Systeme nach Prinzipien der biologischen Evolution*, Frommann–Holzboog, Stuttgart, 1973.

[7] H.-P. Schwefel, *Numerische Optimierung von Computer–Modellen mittels der Evolutionsstrategie*, Birkhäuser, Basel, 1977.

[8] ——, *Numerical Optimization of Computer Models*, Wiley & Sons, Chichester, 1981.