Asymptotic convergence of a simulated annealing algorithm for multiobjective optimization problems

Mario Villalobos-Arias[1], Carlos A. Coello Coello[2*]
and Onésimo Hernández-Lerma[1]
[1]Departamento de Matemáticas
CINVESTAV-IPN, Apartado Postal 14-740
México D.F. 07000, MEXICO
email: {mava,ohernand}@math.cinvestav.mx

[2]Sección de Computación
Depto. de Ingeniería Eléctrica
CINVESTAV-IPN, Av. IPN No. 2508
Col. San Pedro Zacatenco
México, D. F. 07300, MEXICO
email: ccoello@cs.cinvestav.mx

March 12, 2006

Abstract

In this paper we consider a simulated annealing algorithm for multiobjective optimization problems. With a suitable choice of the acceptance probabilities, the algorithm is shown to converge asymptotically, that is, the Markov chain that describes the algorithm converges with probability one to the Pareto optimal set.

Mathematics Subject Classification : 90C29, 90C59, 68T20

Key Words: simulated annealing, multiple objective programming, multiobjective optimization, vector optimization, convergence.

*Corresponding author. Tel. +52 55 5061 3800 x 6564. Fax +52 55 5061 3757

1 Introduction

This paper concerns multiobjective optimization problems in which one wishes to optimize a vector function, say $F(x) = (f_1(x), ..., f_d(x))$.

A typical way to approach these problems is to transform the multiobjective optimization problem into a series of single-objective (or "scalar") problems. This approach indeed makes sense if the functions $f_1, ..., f_d$ are of the same type, but otherwise (for instance, if f_1 denotes distance, f_2 denotes time, and so on) the scalarized problem might be meaningless.

For such cases, one can try a direct approach to the multiobjective optimization problem such as evolutionary algorithms, simulated annealing, and any other related heuristics [2].

Although there have been some convergence proofs for multiobjective evolutionary algorithms (see [10, 11]), most heuristics used for multiobjective optimization do not have such convergence proof reported in the literature. This paper intends to bridge this gap for a class of simulated annealing algorithms.

Here, we consider a simulated annealing algorithm (SAA) for solving multiobjective optimization problems. Under mild assumptions and a suitable choice of the acceptance probabilities, our SAA is shown to converge asymptotically (with probability one) to the Pareto optimal set of the problem.

The remainder of this paper is organized as follows. The multiobjective optimization problem (MOP) is stated in Section 2. In Section 3 we introduce the SAA we are concerned with; we also briefly discuss the algorithm's acceptance probabilities, which are crucial for proving asymptotic convergence. Our main result is stated in Section 4. Finally, our conclusions are provided in Section 5 with some general remarks.

2 The multiobjective optimization problem

To compare vectors in $\mathbb{R}^d$ we will use the standard Pareto order on $\mathbb{R}^d$ defined as follows.

If $\vec{u} = (u_1, u_2, \ldots, u_d)$ and $\vec{v} = (v_1, v_2, \ldots, v_d)$ are vectors in $\mathbb{R}^d$, then

$$\vec{u} \preceq \vec{v} \iff u_i \leq v_i \ \forall \ i \in 1, \ldots, d.$$

This relation is a partial order because it is reflexive, antisymmetric and transitive. We also have $\vec{u} \prec \vec{v} \iff \vec{u} \preceq \vec{v}$ and $\vec{u} \neq \vec{v}$.

The multiobjective optimization problem (MOP) we are concerned with is to find a vector $\vec{x}^* \in X \subset \mathbb{R}^m$ such that

$$F(\vec{x}^*) = \min_{\vec{x} \in X} F(\vec{x}) = \min_{\vec{x} \in X} [f_1(\vec{x}), \ldots, f_n(\vec{x})], \tag{1}$$

where $F : X \subset \mathbb{R}^m \to \mathbb{R}^d$ is a given vector function with components $f_i : X \subset \mathbb{R}^m \to \mathbb{R}$ for each $i \in \{1, \ldots, d\}$, and the minimum is understood in the sense of the Pareto order.

Definition 1:(Pareto optimality).

A vector $\vec{x}^* \in X$ is called a *Pareto optimal solution* for the MOP (1) if

$$\text{there is no } \vec{x} \in X \text{ such that } F(\vec{x}) \prec F(\vec{x}^*).$$

The set $\mathcal{P}^* = \{\vec{x} \in X \ : \ \vec{x} \text{ is a Pareto optimal solution}\}$ is called the *Pareto optimal set* and its image under F, i.e.

$$F(\mathcal{P}^*) \ := \ \{F(\vec{x}) \ : \ \vec{x} \in \mathcal{P}^*\},$$

is the *Pareto front.*

In the sequel we will use the following well–known "scalarization" result.

Proposition 1:

Let $\vec{x}^* \in X$ be a solution of the weighted problem:

$$\min_{\vec{x} \in X} \sum_{s=1}^{d} w_s f_s(\vec{x}), \text{ where } w_s > 0 \ \forall s \in \{1, \ldots, n\} \text{ and } \sum_{s=1}^{d} w_s = 1.$$

Then $\vec{x}^* \in \mathcal{P}^*$.

We omit the proof of this lemma because it is trivial.

As we are concerned with computational aspects, in the remainder of the paper we will replace the set X in (1) with a *finite* set $S \subset I\!R^m$.

3 The simulated annealing algorithm

Nicholas Metropolis et al. [8] originally proposed (in 1953) an algorithm to simulate the evolution of a solid in a heat bath until it reached its thermal equilibrium. The process started from a certain thermodynamic state of the system, defined by a certain energy and temperature. Then, the state was slightly perturbed. If the change in energy produced by this perturbation was negative, the new configuration was accepted. If it was positive, it was accepted with a certain probability. This process was repeated until a frozen state was achieved [4, 12].

About thirty years after the publication of Metropolis' approach, Kirkpatrick et al. [5] and Černy [9] independently pointed out the analogy between this "annealing" process and combinatorial optimization. Such analogies led to the development of an algorithm called "Simulated Annealing" which is a heuristic search technique that has been quite successful in combinatorial optimization problems (see [1] and [6] for details).

The SAA generates a succession of possible solutions of the optimization problem. These possible solutions are the states of a Markov chain and the "energy" of a state is the evaluation of the possible solution that it represents.

The temperature is simulated with a sequence of positive control parameters c_k.

A transition of the Markov chain occurs in two steps, given the value c_k of the control parameter. First, if the current state is i, a new state j is generated with a certain probability $G_{ij}(c_k)$, defined below. Then an "acceptance rule" $A_{ij}(c_k)$ is applied to j. Our main result hinges on a suitable selection of the acceptance rule, which we now discuss.

The generation probability. For each state i, let S_i be a subset of $S \setminus \{i\}$ called the neighborhood of i. We shall assume that the number of elements in S_i is the same, say Θ, for all $i \in S$, and also that the neighbor relation is symmetric, that is, $j \in S_i$ if and only if $i \in S_j$. Then, denoting by χ_{S_i} the indicator function of S_i (i.e. $\chi_{S_i}(j) := 1$ if $j \in S_i$ and 0 otherwise), we define the generation probability

$$G_{ij}(c_k) := \frac{\chi_{S_i}(j)}{\Theta} \quad \text{for all } i, j \in S. \tag{2}$$

The acceptance probability. The acceptance probability, which is crucial for the behavior of the SAA, can be defined in several different ways. For instance, Serafini [13] proposes to use the L_∞–Tchebycheff norm given by

$$A'_{ij}(c) = \min\left\{1, \exp\left(\max_{s\in\{1,\ldots,d\}} \frac{\lambda_s(f_s(i) - f_s(j))}{c}\right)\right\},$$

where the λ_s are given positive parameters. This acceptance probability has the "inconvenience" that if a single entry is improved (i.e. $f_s(i) > f_s(j)$ for some s) or has the same value, then the state j is accepted, which obviously is not very good. For example, in Figure 1, in which $f_1(j) = f_1(i)$, we have $A_{ij} = 1$ although $f_2(j)$ is too "bad" in comparison with $f_2(i)$.

On the other hand, Ulungu and coworkers [14, 15, 17, 16] use

$$\begin{aligned} A''_{ij}(c) &:= \min\left\{1, \exp\left(\sum_{s=1}^{d} \frac{\lambda_s(f_s(i) - f_s(j))}{c}\right)\right\} \\ &= \exp\left\{-\left(\sum_{s=1}^{d} \frac{\lambda_s(f_s(j) - f_s(i))}{c}\right)^+\right\}, \end{aligned} \tag{3}$$

where as usual, let a^+ be the positive part of a number $a \in \mathbb{R}$, namely

$$a^+ := \begin{cases} a & \text{if } a > 0, \\ 0 & \text{otherwise.} \end{cases}$$

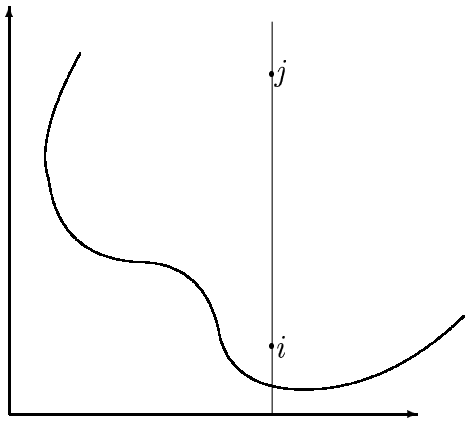

Figure 1: Graphical illustration of the "inconvenience" of the acceptance probability $A'_{ij}(c)$ proposed by Serafini [13].

But again, in this case it is possible to have the acceptance probability depending on the change of a *single* entry $f_s(i) - f_s(j),\ s = 1, \ldots, n$.

Here, we shall use the acceptance probability [13]

$$A_{ij}(c) := \prod_{s=1}^{d} \min\left\{1, \exp\left(\frac{f_s(i) - f_s(j)}{c}\right)\right\},$$

which can be expressed in the simpler form

$$A_{ij}(c) = \exp\left(-\frac{\sum_{s=1}^{d}(f_s(j) - f_s(i))^+}{c}\right). \tag{4}$$

This acceptance probability is obviously "better" than the one in (3) because only the entries that do not improve are taken into account to calculate the probability; this probability could be improved changing c by an individual c_s for each entry $s = 1, \ldots, n$.

The transition probability. Having the generation and the acceptance probabilities, we can now define the *transition probability* from i to j as

$$P_{ij}(c_k) := \begin{cases} G_{ij}(c_k)A_{ij}(c_k) & \text{if } i \neq j, \\ 1 - \sum_{l \in S, l \neq i} P_{il}(c_k) & \text{if } i = j \end{cases} \tag{5}$$

where A_{ij} is as in (4) (or as in (3)).

Note that for theoretical purposes we can use $f_s(i) - f_s(j)$ instead of $\lambda_s(f_s(i) - f_s(j))$ or $(f_s(i) - f_s(j))/c_s$, because the last two expressions can be transformed into the first one via the changes $g_s = \lambda_s f_s$ or $g_s = f_s/c_s$, respectively. Hence, at the remainder of this work we will use the first one.

4 Main result

Let

$$\Sigma_{opt} := \{i \in S \ : \ \sum_{s=1}^{d} f_s(i) = \Sigma_m\},$$

where

$$\Sigma_m := \min_{j \in S} \sum_{s=1}^{d} f_s(j). \tag{6}$$

Then, by Proposition 1, the Pareto optimal set $\mathcal{P}^*$ contains Σ_{opt}, i.e.

$$\Sigma_{opt} \subset \mathcal{P}^*. \tag{7}$$

We next present our main result, which in particular states the convergence of the SAA for the MOP (1). The convergence is understood in the following sense.

Definition 2:
Let $P(c)$ be the transition matrix associated with the SAA defined by (2), (4), (5), and let $\{X_k(c), k = 0, 1, 2, ...\}$ be the corresponding Markov chain, at temperature c. The SAA is said to *converge with probability 1* if

$$\lim_{c \searrow 0} \lim_{k \to \infty} I\!P\{X_k(c) \in \mathcal{P}^*\} = 1.$$

The next theorem, which is the main result in this paper, is an extension to MOPs of the results presented in [1]. Here we use ideas similar to those in that paper, with the appropriate changes.

In the proof of this theorem we show that the algorithm converges to the set $\Sigma_{opt} \subseteq \mathcal{P}^*$, because of the particular transition probability we use.

Theorem 1:
Let $P(c)$ be as in Definition 2 and, moreover, suppose that $G(c)$ is irreducible. Then:

(a) The Markov chain has a stationary distribution $\vec{q}(c)$, whose components are given by

$$q_i(c) = \frac{1}{N_0(c)} \exp\left(-\frac{\sum_{s=1}^{d} f_s(i)}{c}\right) \quad \forall\ i \in S, \tag{8}$$

where

$$N_0(c) = \sum_{j \in S} \exp\left(-\frac{\sum_{s=1}^{d} f_s(j)}{c}\right) \tag{9}$$

(b) For each $i \in S$

$$q_i^* := \lim_{c \searrow 0} q_i(c) = \frac{1}{|\Sigma_{opt}|} \chi_{\Sigma_{opt}}(i),$$

where $|\Sigma_{opt}|$ denotes the number of elements in Σ_{opt}.

(c) The SAA converges with probability 1.

These results remain valid if (4) is replaced with (3).

Before presenting the proof of Theorem 1 we state some preliminary results. First, we note the following fact, which is due to $a^+ = a + (-a)^+$ $(= a + a^-)$.

Lemma 1:

For any real numbers $a_1, a_2, \ldots, a_n,\ b_1, b_2, \ldots, b_n$,

$$\begin{aligned} \sum_{k=1}^{d}(a_k - b_k) + \left(\sum_{k=1}^{d}(b_k - a_k)\right)^+ &= \left(\sum_{k=1}^{d}(a_k - b_k)\right)^+, \\ \sum_{k=1}^{d}(a_k - b_k) + \sum_{k=1}^{d}(b_k - a_k)^+ &= \sum_{k=1}^{d}(a_k - b_k)^+. \end{aligned}$$

We will need some properties of the limiting distribution, which we will present next. Recall that a probability distribution $\vec{q}$ is called the *limiting distribution* of a Markov chain with transition probability P if

$$q_i = \lim_{k \to \infty} I\!P(X_k = i | X_0 = j) \text{ for all } i, j \in S.$$

If such a limiting distribution $\vec{q}$ exists and $a_i(k) = I\!P(X_k = i)$, for $i \in S$, denotes the distribution of X_k, then

$$\lim_{k \to \infty} a_i(k) = q_i \text{ for all } i \in S.$$

Moreover, $\vec{q}$ is an *invariant* (or *stationary*) distribution of the Markov chain, which means that

$$\vec{q} = \vec{q}\,P; \tag{10}$$

that is, $\vec{q}$ is a left eigenvector of P with eigenvalue 1. A converse to this result (which is true for *finite* Markov chains) is given in Lemma 3 below.

Observe that (10) trivially holds if $\vec{q}$ is a probability distribution satisfying

$$q_i P_{ij} = q_j P_{ji} \quad \forall i, j \in S. \tag{11}$$

Equation (11) is called the *detailed balance equation*, and (10) is called the *global balance equation*.

It is well known that in an irreducible Markov chain all of the states have the same period. This observation yields the following.

Lemma 2:

An irreducible Markov chain with transition matrix P is aperiodic if there exists $j \in S$ such that $P_{jj} > 0$.

Lemma 3: [7, pag.19]

Let P be the transition matrix of a finite, irreducible and aperiodic Markov chain. Then the chain has a unique stationary distribution $\vec{q}$, that is $\vec{q}$ is the unique distribution that satisfies (10), and, in addition, $\vec{q}$ is the chain's limiting distribution

Proof of Theorem 1

(a) Since G is irreducible, using Lemma 2 it can be seen that the Markov chain is irreducible and aperiodic (see [1, p.39]). Hence, by Lemma 3 there exists a unique stationary distribution. We now use (2) and (5) to see that (11) holds for all $i \neq j$. First note that

$$\begin{aligned} q_i(c)P_{ij}(c) &= q_i(c)G_{ij}(c)A_{ij}(c) \\ &= \begin{cases} \frac{1}{\Theta} q_i(c)A_{ij}(c) & \text{if } j \in S_i \\ 0 & \text{if } j \notin S_i. \end{cases} \end{aligned}$$

Similarly,

$$\begin{aligned} q_j(c)P_{ji}(c) &= q_j(c)G_{ji}(c)A_{ji}(c) \\ &= \begin{cases} \frac{1}{\Theta} q_j(c)A_{ji}(c) & \text{if } i \in S_j \\ 0 & \text{if } i \notin S_j. \end{cases} \end{aligned}$$

Thus, since $i \in S_j$ if and only if $j \in S_i$, to obtain (11) we only have to prove that

$$q_i(c)A_{ij}(c) = q_j A_{ji}(c).$$

But this follows from (4), (8) and Lemma 1, because

$$
\begin{aligned}
q_i(c)A_{ij}(c) \;&= \\
&= \frac{1}{N_0(c)} \exp\left(-\frac{\sum_{s=1}^{d} f_s(i)}{c}\right) \exp\left(-\frac{\sum_{s=1}^{d} (f_s(j) - f_s(i))^+}{c}\right) \\
&= \frac{1}{N_0(c)} \exp\left(-\frac{\sum_{s=1}^{d} f_s(j)}{c}\right) \exp\left(-\frac{\sum_{s=1}^{d} (f_s(i) - f_s(j)) + \sum_{s=1}^{d} (f_s(j) - f_s(i))^+}{c}\right) \\
&= \frac{1}{N_0(c)} \exp\left(-\frac{\sum_{s=1}^{d} f_s(j)}{c}\right) \exp\left(-\frac{\sum_{s=1}^{d} (f_s(i) - f_s(j))^+}{c}\right) \\
&= q_j(c)A_{ji}(c).
\end{aligned}
$$

This shows that (11) holds, which in turn yields part (a) in Theorem 1. (Note that this proof, with obvious changes, remains valid if the acceptance probability is given by (3) rather than (4)).

(b) Note that for each $a \leq 0$

$$
\lim_{x \searrow 0} e^{a/x} = \begin{cases} 1 & \text{if } a = 0, \\ 0 & \text{otherwise.} \end{cases} \tag{12}
$$

Now, by (6), (8) and (9)

$$
\begin{aligned}
q_i(c) &= \frac{\exp\left(-\frac{\sum_{s=1}^{d} f_s(i)}{c}\right)}{\sum_{j\in S}\exp\left(-\frac{\sum_{s=1}^{d} f_s(j)}{c}\right)} \\
&= \frac{\exp\left(\frac{\Sigma_m-\sum_{s=1}^{d} f_s(i)}{c}\right)}{\sum_{j\in S}\exp\left(\frac{\Sigma_m-\sum_{s=1}^{d} f_s(j)}{c}\right)} \\
&= \frac{\exp\left(\frac{\Sigma_m-\sum_{s=1}^{d} f_s(i)}{c}\right)}{\sum_{j\in S}\exp\left(\frac{\Sigma_m-\sum_{s=1}^{d} f_s(j)}{c}\right)}\left(\chi_{\Sigma_{opt}}(i)+\chi_{S-\Sigma_{opt}}(i)\right) \\
&= \frac{1}{\sum_{j\in S}\exp\left(\frac{\Sigma_m-\sum_{s=1}^{d} f_s(j)}{c}\right)}\chi_{\Sigma_{opt}}(i) \\
&\quad + \frac{\exp\left(\frac{\Sigma_m-\sum_{s=1}^{d} f_s(i)}{c}\right)}{\sum_{j\in S}\exp\left(\frac{\Sigma_m-\sum_{s=1}^{d} f_s(j)}{c}\right)}\chi_{S-\Sigma_{opt}}(i).
\end{aligned}
$$

Now let $c \searrow 0$. Then, by (12), the second term of the latter sum goes to 0, whereas the denominator of the first term goes to $|\Sigma_{opt}|$. Hence

$$\lim_{c\searrow 0} q_i(c) = \frac{1}{|\Sigma_{opt}|}\chi_{\Sigma_{opt}}(i) \;+\; 0 \;=\; q_i^*,$$

which completes the proof of part (b).

(c) By (b) and Lemma 3

$$\lim_{c\searrow 0}\lim_{k\to\infty} I\!P\{X_k = i\} = \lim_{c\searrow 0} q_i(c) = q_i^*,$$

and so by (7)

$$\lim_{c\searrow 0}\lim_{k\to\infty} I\!P\{X_k \in \mathcal{P}^*\} \geq \lim_{c\searrow 0}\lim_{k\to\infty} I\!P\{X_k \in \Sigma_{opt}\} = 1. \tag{13}$$

Thus

$$\lim_{c\searrow 0}\lim_{k\to\infty} I\!P\{X_k \in \mathcal{P}^*\} = 1,$$

and (c) follows. ■

5 Concluding remarks

We have shown in Theorem 1 that a suitable choice of the acceptance probabilities yields the asymptotic convergence of the SAA. This is reassuring, of course, because it means that the algorithm is indeed heading in the right direction. However, for computational purposes, our approach might not be very useful.

Indeed, what we actually prove is that, as in (13), the underlying Markov chain converges to the set Σ_{opt} which can be very "small" compared to the Pareto optimal set $\mathcal{P}^*$.

This is illustrated in Figure 2 in which the Pareto front corresponds to the parts on the boundary of S joining the points A and B, and also the points C and D, whereas $F(\Sigma_{opt})$ corresponds only to the points that give p_1 and p_2.

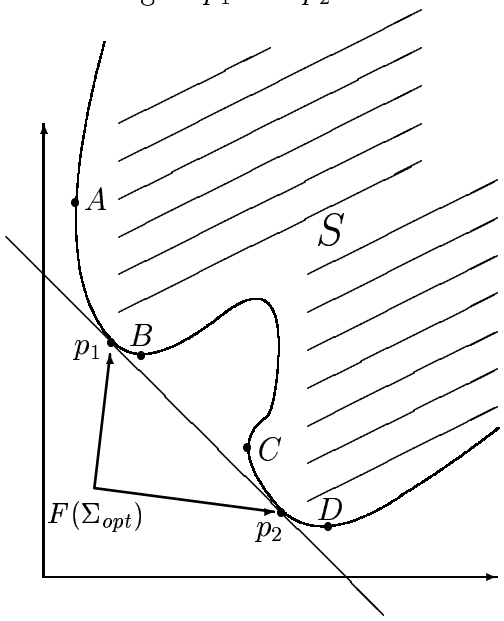

Figure 2: Comparison of Σ_{opt} and $\mathcal{P}^*$

To improve our SAA one possibility would be to introduce an "elite set", which is a standard procedure in multiobjective evolutionary algorithms [2, 3]. At each step of the algorithm, the elite set contains all the nondominated points generated so far. Thus, by introducing the elite set, the idea would be to make the contents of such elite set to converge to the Pareto optimal set. Research along these lines is in progress.

Acknowledgements

The second author acknowledges support fom CONACyT through project 42435-Y. The third author was partially supported by CONACyT Grant 45693.

References

[1] E.H. Aarts and J.H. Korst. *Simulated Annealing and Boltzmann Machines: A Stochastic Approach to Combinatorial Optimization and Neural Computing.* John Wiley & Sons, Chichester, UK, 1989.

[2] Carlos A. Coello Coello, David A. Van Veldhuizen, and Gary B. Lamont. *Evolutionary Algorithms for Solving Multi-Objective Problems.* Kluwer Academic Publishers, New York, May 2002. ISBN 0-3064-6762-3.

[3] Kalyanmoy Deb. *Multi-Objective Optimization using Evolutionary Algorithms.* John Wiley & Sons, Chichester, UK, 2001. ISBN 0-471-87339-X.

[4] Kathryn A. Dowsland. Simulated Annealing. In Colin R. Reeves, editor, *Modern Heuristic Techniques for Combinatorial Problems*, chapter 2, pages 20–69. John Wiley & Sons, 1993.

[5] S. Kirkpatrick, C.D. Gellatt, and M.P. Vecchi. Optimization by Simulated Annealing. *Science*, 220(4598):671–680, 1983.

[6] P. Laarhoven and E.H. Aarts. *Simulated Annealing: Theory and Applications.* D. Reidel, Boston, Massachusetts, 1987.

[7] G.F. Lawler. *Introduction to Stochastic Processes.* Chapman & Hall/CRC, Boca Raton, Florida, 1995.

[8] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A. Teller, and E. Teller. Equation of State Calculations by Fast Computing Machines. *Journal of Chemical Physics*, 21(6):1087–1092, 1953.

[9] V. Černy. A Thermodynamical Approach to the Traveling Salesman Problem: An Efficient Simulation Algorithm. *Journal of Optimization Theory and Applications*, 45(1):41–51, 1985.

[10] Günter Rudolph. On a Multi-Objective Evolutionary Algorithm and Its Convergence to the Pareto Set. In *Proceedings of the 5th IEEE Conference on Evolutionary Computation*, pages 511–516, Piscataway, New Jersey, 1998. IEEE Press.

[11] Günter Rudolph and Alexandru Agapie. Convergence Properties of Some Multi-Objective Evolutionary Algorithms. In *Proceedings of the 2000 Conference on Evolutionary Computation*, volume 2, pages 1010–1016, Piscataway, New Jersey, July 2000. IEEE Press.

[12] Sadiq M. Sait and Habib Youssef. *Iterative Computer Algorithms with Applications in Engineering. Solving Combinatorial Optimization Problems.* IEEE Computer Society, Los Alamitos, California, 1999.

[13] Paolo Serafini. Simulated Annealing for Multiple Objective Optimization Problems. In G.H. Tzeng, H.F. Wang, U.P. Wen, and P.L. Yu, editors, *Proceedings of the Tenth International Conference on Multiple Criteria Decision Making: Expand and Enrich the Domains of Thinking and Application*, volume 1, pages 283–292, Berlin, 1994. Springer-Verlag.

[14] E.L. Ulungu. *Optimisation Combinatoire multicritere: Determination de l'ensemble des solutions efficaces et methodes interactives.* PhD thesis, Faculté des Sciences, Université de Mons-Hainaut, Mons, Belgium, 1993.

[15] E.L. Ulungu, J. Teghem, and Ph. Fortemps. Heuristics for multi-objective combinatorial optimization by simulated annealing. In J. Gu, G. Chen, Q. Wei, and S. Wang, editors, *Multiple Criteria Decision Making: Theory and Applications. Proceedings of the 6th National Conference on Multiple Criteria Decision Making*, pages 228–238, Windsor, UK, 1995. Sci-Tech.

[16] E.L. Ulungu, J. Teghem, Ph. Fortemps, and D. Tuyttens. MOSA Method: A Tool for Solving Multiobjective Combinatorial Optimization Problems. *Journal of Multi-Criteria Decision Analysis*, 8(4):221–236, 1999.

[17] E.L. Ulungu, J. Teghem, and Ch. Ost. Efficiency of interactive multi-objective simulated annealing through a case study. *Journal of the Operational Research Society*, 49:1044–1050, 1998.