Evolutionary Search under Partially Ordered Fitness Sets

Günter Rudolph*

March 16, 1999

Abstract

The search for minimal elements in partially ordered sets is a generalization of the task of finding Pareto-optimal elements in multi-criteria optimization problems. Since there are usually many minimal elements within a partially ordered set, a population-based evolutionary search is, as a matter of principle, capable of finding several minimal elements simultaneously and gains therefore a steadily increase of popularity. Here, we present an evolutionary algorithm which population converges with probability one to the set of minimal elements within a finite number of iterations.

1 Introduction

The search for minimal elements in partially ordered sets is a generalization of the task of finding Pareto-optimal elements in multi-criteria optimization problems. Since there are usually many minimal elements within a partially ordered set, a population-based evolutionary search is, as a matter of principle, capable of finding several minimal elements simultaneously and gains therefore a steadily increase of popularity. This increase of popularity is witnessed by numerous proposals of multi-criteria evolutionary algorithms during the last few years – this rapid development was, however, not accompanied by a comparable build-up of a theoretical foundation.
But the first steps towards an elimination of this shortcoming has been made: It was shown in [1] in case of finite search sets that an evolutionary algorithm (EA) with 'positive variation kernel' and 'elite preservation strategy' (these notions are explained later) is capable of generating a sequence of populations such that at least one individual enters the set of minimal elements of the partially ordered fitness set in finite time with probability one and stays there forever. Moreover, it was proven that the population of such an EA converges completely to the set of minimal elements if the population at step $t+1$ is just the set of minimal elements of the union of the population of parents and the generated offspring at step t. Evidently, the population size is not fixed in this case; it grows to the size of the set of minimal elements which may be prohibitively large. Therefore, it is the goal of this paper to devise an evolutionary algorithm with *fixed* population size which population converges to the set of minimal elements. Such an EA is described in Section 3 and analyzed in Section 4. Basic terminology is introduced in Section 2.

*Department of Computer Science, University of Dortmund, Germany.

2 Partially Ordered Sets

A prerequisite to introduce 'partially ordered sets' is the notion of the 'relation.' The definitions presented in this section are extracted from Ester [2] and Trotter [3].

Definition 1 *Let $\mathcal{X}$ be some set. The subset $\mathcal{R} \subseteq \mathcal{X} \times \mathcal{X}$ is called a* binary relation *in $\mathcal{X}$. Let $x, y \in \mathcal{X}$. If $(x, y) \in \mathcal{R}$, also denoted $x\mathcal{R}y$, then x is said to be in relation $\mathcal{R}$ to y. A relation $\mathcal{R}$ in $\mathcal{X}$ is said to be*

(a) reflexive *if $x\mathcal{R}x$ is true for all $x \in \mathcal{X}$,*
(b) antireflexive *if $x\mathcal{R}y \Rightarrow x \neq y$ is true for all $x, y \in \mathcal{X}$,*
(c) symmetric *if $x\mathcal{R}y \Rightarrow y\mathcal{R}x$ is true for all $x, y \in \mathcal{X}$,*
(d) antisymmetric *if $x\mathcal{R}y \wedge y\mathcal{R}x \Rightarrow x = y$ is true for all $x, y \in \mathcal{X}$,*
(e) asymmetric *if $x\mathcal{R}y \Rightarrow \overline{y\mathcal{R}x}$ is true for all $x, y \in \mathcal{X}$,*
(f) transitive *if $x\mathcal{R}y \wedge y\mathcal{R}z \Rightarrow x\mathcal{R}z$ is true for all $x, y, z \in \mathcal{X}$.* □

Some relations that possess several of the properties above simultaneously bear their own names. For example, if $\mathcal{R}$ is a reflexive, symmetric, and transitive relation then $\mathcal{R}$ is called an *equivalence relation*. In this case it is common to use the symbol "$\sim$" in lieu of $\mathcal{R}$. A reflexive, antisymmetric, and transitive relation "$\preceq$" is termed a *partial order relation* whereas a *strict partial order relation* "$\prec$" must be antireflexive, asymmetric, and transitive. The latter relation may be obtained by the former one by setting $x \prec y := (x \preceq y) \wedge (x \neq y)$. After these preparations one is in the position to turn to the actual objects of interest.

Definition 2 *Let $\mathcal{X}$ be some set. If the partial order relation "$\preceq$" is valid on $\mathcal{X}$ then the pair $(\mathcal{X}, \preceq)$ is called a* partially ordered set *(or short:* poset*). If $x \prec y$ for some $x, y \in \mathcal{X}$ then x is said to* dominate *y. Distinct points $x, y \in \mathcal{X}$ are said to be* comparable *when either $x \prec y$ or $y \prec x$. Otherwise, x and y are* incomparable *which is denoted by $x \parallel y$. If each pair of distinct points of a poset $(\mathcal{X}, \preceq)$ is comparable then $(\mathcal{X}, \preceq)$ is called a* totally ordered set *or a* chain*. Dually, if each pair of distinct points of a poset $(\mathcal{X}, \preceq)$ are incomparable then $(\mathcal{X}, \preceq)$ is termed an* antichain*.* □

For example, $(\mathbb{R}^n, \preceq)$ with $n \geq 2$ is a partially ordered set when $x \preceq y$ means $x_i \leq y_i$ for all $i = 1, \ldots, n$. One obtains a strict partial order relation "$\prec$" from this partial order relation if it is additionally required that $x \neq y$. Notice that the poset $(\mathbb{R}^n, \preceq)$ is neither a chain nor an antichain. The situation changes for the poset $(\mathbb{R}, \preceq)$ with $x \preceq y$ if and only if $x \leq y$. Since each pair of distinct points in $\mathbb{R}$ is comparable the poset $(\mathbb{R}, \preceq)$ is totally ordered and therefore a chain. An example for an antichain is the set of "minimal elements" introduced next.

Definition 3 *An element $x^* \in \mathcal{X}$ is called a* minimal element *of the poset $(\mathcal{X}, \preceq)$ if there is no $x \in \mathcal{X}$ such that $x \prec x^*$. The set of all minimal elements, denoted $\mathcal{M}(\mathcal{X}, \preceq)$, is said to be* complete *if for each $x \in \mathcal{X}$ there is at least one $x^* \in \mathcal{M}(\mathcal{X}, \preceq)$ such that $x^* \preceq x$.* □

Minimal elements are the targets of the evolutionary search studied here. Since the analysis presented shortly requires the completeness of $\mathcal{M}(\mathcal{X}, \preceq)$ it is useful to know under which circumstances this assumption is fulfilled. If the poset $(\mathcal{X}, \preceq)$ is finite then the completeness of $\mathcal{M}(\mathcal{X}, \preceq)$ is guaranteed ([4], p. 91). This result shows that the set of minimal elements may be incomplete only if the poset is infinitely large. Sufficient conditions for the completeness of $\mathcal{M}(\mathcal{X}, \preceq)$ in case of infinitely large posets $(\mathcal{X}, \preceq)$ may be found, for example, in [5]. This general case is beyond the scope of this paper – hereinafter it is assumed that the posets are always finite and hence endowed with a complete set of minimal elements.

3 Evolutionary Algorithm

Let S be some finite search set and $f : S \to \mathcal{F} = \{f(x) : x \in S\}$ the fitness function with partially ordered fitness values, i.e., $(\mathcal{F}, \preceq)$ is a poset. An individual of the evolutionary algorithm is represented by the pair $(x, \psi) \in S \times \Psi$ where Ψ is a compact subset of $\mathbb{R}^m$. Here, ψ represents the values of m parameters that may affect, for example, the mutation distribution or any other procedure that is involved in the production of offspring. The mapping $f : S \to \mathcal{F}$ induces also a partial order relation "$\preceq_f$" on the search set S (and similarly on the set of individuals) via the definitions

$$\begin{array}{lllllllll} x_1 & \prec_f & x_2 & \Leftrightarrow & f(x_1) & \prec & f(x_2) & & \\ x_1 & \sim_f & x_2 & \Leftrightarrow & f(x_1) & = & f(x_2) & & \\ x_1 & \preceq_f & x_2 & \Leftrightarrow & x_1 & \prec_f & x_2 & \vee & x_1 \sim_f x_2 \,. \end{array}$$

For the sake of notational convenience, the subscript f will be omitted hereinafter, i.e., the statement $x_1 \preceq x_2$ will actually mean $x_1 \preceq_f x_2$, and an analogous convention applies to the remaining relations. The targets of the evolutionary search are the elements of $\mathcal{M}(\mathcal{F}, \preceq)$. Clearly, whenever the fitness value $f(x)$ of an individual (x, ψ) is an minimal element of the poset $(\mathcal{F}, \preceq)$ then x is a minimal element of the poset $(S, \preceq_f)$ and vice versa.
The pseudo code of the evolutionary algorithm considered here is presented in Fig. 1. At the beginning, μ individuals are initialized arbitrarily from the set $S \times \Psi$. This yields the population P_0. After setting the generation counter to $t = 0$ the EA enters the loop in which each iteration represents the production and selection process of one generation. Each iteration can be divided into three phases.

Phase 1: At first, μ parents of the current population P_t produce λ offspring in some probabilistic manner ($\lambda \geq \mu \geq 1$). The offspring are collected in the multi-set Q (duplicate members are not discarded). Those offspring which are minimal among all offspring are moved to Q^* and the auxiliary multi-sets P' and Q' are emptied.
At the end of phase 1, the offspring are partitioned into the multi-sets Q and Q^* with $|Q^*| \geq 1$ and $|Q| + |Q^*| = \lambda$. Every offspring in Q is worse than some offspring in Q^*.

Phase 2: For each offspring q from Q^* let $D(q)$ contain all parents from P_t that are dominated by offspring q. If such parents exist then they are discarded from P_t and the offspring q is moved from Q^* to P'. If no parent was dominated but offspring q is incomparable to all parents then q is moved from Q^* to Q'.
At the end of phase 2, set P' contains offspring that are better than some parent, set Q' contains those offspring that are either better than some parent or incomparable to all parents, and Q^* now contains offspring being not better than some parent. Those parents which are left over in P_t are incomparable to each offspring in $P' \cup Q'$. Clearly, every offspring in Q is worse than any offspring in $P' \cup Q' \cup Q^*$.

Phase 3: The multi-set P_{t+1} of parents of the next iteration consists of the union of P' and the residual multi-set P_t. By construction, it is guaranteed[1] that $|P_{t+1}| = |P_t \cup P'| = |P_t| + |P'| \leq \mu$. If $|P_{t+1}| < \mu$ then members of Q' are moved to P_{t+1}. If Q' contains more members than necessary to fill P_{t+1} an arbitrary rule may be applied to choose the members to be moved to P_{t+1}. If Q' contains too few members to fill P_{t+1} the same procedure is applied to Q^* and, if necessary, to Q. Since $\lambda \geq \mu$ it is guaranteed that the new population can be completed to μ members in this manner. If λ is less than μ then P_{t+1} might be filled with randomly generated individuals.
At the end of phase 3, each member of the *original* population P_t (at the beginning of phase 1) which is not dominated by some offspring has been passed to the new population P_{t+1} whereas each dominated parent has been replaced by some better offspring.

4 Analysis

By construction of the algorithm just presented one can easily deduce some auxiliary results that facilitate the proof of the main result. For example, if an optimal individual is already a member of the population P_t then it will be also a member of the next population P_{t+1}. This fact may be formulated as follows:

Lemma 1
Let $x \in \mathcal{M}(S, \preceq)$. If $x \in P_t$ then $x \in P_{t+1}$ for $t \geq 0$. □

Suppose that an optimal offspring has been produced which is not contained in the parent population P_t. Two things may happen. First, the offspring dominates a parent in the current population P_t. In this case it will move to P' and finally to the new population P_{t+1}. Second, there is no parent in the current population P_t that is dominated by the optimal offspring. In this case the offspring will move to Q' and it is not guaranteed that it will also enter the new population P_{t+1}. Thus, an optimal offspring may get lost although there exist (incomparable) parents that are not optimal! This situation is summarized below.

Lemma 2
Let $x \in \mathcal{M}(S, \preceq)$. If $x \in Q$ but $x \notin P_t$ for some $t \geq 0$ then either $x \in P'$ or $x \in Q'$. Moreover, if $x \in P'$ then $x \in P_{t+1}$. □

If all parents are optimal we are done. Suppose there exist parents which are not optimal. Since the set of minimal elements is complete it is guaranteed that there exists a minimal element that dominates a non-optimal parent. In symbols:

Lemma 3
$\exists\, y \in P_t : y \notin \mathcal{M}(S, \preceq) \;\;\Rightarrow\;\; \exists\, x \in \mathcal{M}(S, \preceq) : x \prec y.$ □

Evidently, one needs a mechanism that guarantees the creation of such elements (offspring) since such an event would ensure the assignment of x to P' in Lemma 2. A sufficient criterion for this purpose is a ‘positive variation kernel.’

[1] If some $q \in Q^*$ enters P' in phase 2 then at least one member of P_t is deleted.

```
initialize P_0; set t = 0
repeat
  (* PHASE 1 *)
  Q = offspring(P_t)
  Q* = M(Q, ⪯)
  Q = Q \ Q*
  P' = Q' = ∅
  (* PHASE 2 *)
  for each q ∈ Q*:
    D(q) = {p ∈ P_t : q ≺ p}
    if D(q) ≠ ∅ then
      P_t = P_t \ D(q)
      P' = P' ∪ {q}
      Q* = Q* \ {q}
    endif
    if D(q) = ∅ ∧ q ∥ p for all p ∈ P_t then
      Q' = Q' ∪ {q}
      Q* = Q* \ {q}
    endif
  endfor
  (* PHASE 3 *)
  P_{t+1} = P_t ∪ P'
  if |P_{t+1}| < μ then
    fill P_{t+1} with elements from:
      1. Q'
      2. Q*
      3. Q
    until P_{t+1} = μ
  endif
  t = t + 1
until stopping criterion fulfilled
```

Figure 1: Pseudo code of the evolutionary algorithm with partially ordered fitness.

Definition 4
Let ρ with $1 \leq \rho \leq \mu$ denote the number of parents that participate in the process of producing a single offspring $(y, \psi) \in S \times \Psi$ where S is the search set and Ψ is a fixed compact subset of $\mathbb{R}^m$. A transition probability function $\mathsf{K} : (S \times \Psi)^\rho \times (S \times \Psi) \to [0, 1]$ with the property

$$\mathsf{K}(x_1, \psi_1, x_2, \psi_2, \ldots, x_\rho, \psi_\rho; y, \psi) \geq \delta > 0$$

for all $y, x_1, \ldots, x_\rho \in S$ and $\psi, \psi_1, \ldots, \psi_\rho \in \Psi$ is termed a *positive variation kernel*. □

The positiveness of a variation kernel can be achieved easily. For example, suppose that the search set is the set of binary strings of length ℓ and that a new offspring is produced by (one point or uniform) crossover with crossover probability ψ_1 and the usual bit-flipping mutation, i.e., each bit is inverted independently with mutation probability ψ_2. Even if ψ_1 and ψ_2 are controlled by some exogenous schedule or some self-adapting mechanism, the positiveness of the variation kernel (representing the joint transition probabilities of crossover and mutation) is guaranteed as long as $\psi_2 \in [a, b] \subset \mathbb{R}$ with $0 < a \leq b < 1$. Further examples may be found in [6, 7].

Lemma 4
If T denotes the random number of trials necessary to generate a specific offspring from an arbitrary collection of parents with a positive variation kernel then $\mathsf{P}\{ T < \infty \} = 1$.

Proof:
Since the variation kernel is positive the probability that a specific offspring is not generated from an arbitrary collection of parents within t trial is $\mathsf{P}\{ T > t \} \leq (1-\delta)^t$. As a consequence, one immediately obtains $\mathsf{P}\{ T < \infty \} = 1 - \lim_{t \to \infty} \mathsf{P}\{ T > t \} \geq 1 - \lim_{t \to \infty} (1-\delta)^t = 1$. □

Now one is in the position to prove the main result:

Theorem 1
Let the variation kernel of the evolutionary algorithm described in Figure 1 be positive. Then the population entirely consists of minimal elements after a finite number of iterations with probability one.

Proof:
Suppose that no member of the population at some step $t \geq 0$ is optimal. Lemma 3 ensures that there exist a minimal element that dominates at least one of the parents. Owing to Lemma 4 this minimal element can be produced by the variation operators in a finite number of steps with probability 1. It follows from Lemma 2 that this optimal offspring will move to P' and finally to P_{t+1}. Lemma 1 guarantees that this optimal individual will stay in the population forever. A μ-fold repetition of this argumentation leads to the conclusion that the entire population of the evolutionary algorithm consists of minimal elements after a finite number of steps with probability one. □

It should be mentioned that the evolutionary algorithm considered here realizes a stronger version of the 'elite preservation strategy' than introduced in [1]: Unless there is an offspring that dominates a specific parent, this parent will also be a parent of the next iteration. This stronger version is apparently necessary for proving the convergence of the entire population to the set of minimal elements.
An example of an evolutionary algorithm that violates elite preservation is as follows: Suppose that μ parents produce λ offspring with a positive variation kernel. Let M be the set of minimal elements relative to the union of parents and offspring. In the algorithm of Peschel & Riedel [8], the set M is exactly the population of parents of the next iteration. Needless to say, in this case the size of the population is not constant over time and it will finally grow to the cardinality of the set of minimal elements [1]. This kind of selection was later re-invented by several authors—with the difference that the population size was kept fixed. This property is usually achieved by adding some individuals if the size of M is less than μ and by deleting some individuals from M at random if the size of M is larger than μ.
This method does not lead to convergence: Suppose that all μ parents at iteration $t \geq 0$ represent minimal elements and that the cardinality of the set of minimal elements is at least larger than $2\,\mu$. Moreover, $\lambda = \mu$. Since the variation kernel is positive there exists a positive minimum probability

that the λ offspring are not minimal elements and that parents as well as offspring are mutually incomparable. Since $|M| = 6 > \mu = 3$, three members of M are deleted at random. With probability $3!/(6 \cdot 5 \cdot 4) = 1/20$ all optimal parents will be removed from M such that the population of parents of the next iteration will not contain any minimal element. Thus, optimal individuals will be found and lost, found and lost and so forth with some minimum probability. Clearly, such a behavior precludes the property of convergence. But there is a simple remedy: If the cardinality of M is larger than μ then one should delete only those members of M at random which were not parents. In this case the 'strong elite preservation property' is not violated and one obtains convergence of the entire population to the set of minimal elements.

5 Conclusions

An evolutionary algorithm which population is guaranteed to converge to the set of minimal elements in a finite number of iterations has been proposed. The more important contribution of this work, however, is the observation which properties of the evolutionary algorithm are sufficient to prove the convergence. These properties are (i) a positive variation kernel and (ii) the strong elite preservation strategy. Future work should therefore be engaged in examining other evolutionary algorithms with respect to these properties. Since these (sufficient) conditions were only proved for finite search sets a generalization to infinite search sets is desirable. Some work on such search sets is available [9, 10] albeit specialized to multi-criteria problems. It would be instructive to generalize these results to the problem of finding minimal elements of arbitrary partially ordered sets.

Acknowledgments

This work is a result of the *Collaborative Research Center "Computational Intelligence" (SFB 531)* supported by the Deutsche Forschungsgemeinschaft (DFG).

References

[1] G. Rudolph. Evolutionary search for minimal elements in partially ordered finite sets. In V. W. Porto, N. Saravanan, D. Waagen, and A. E. Eiben, editors, *Evolutionary Programming VII, Proceedings of the 7th Annual Conference on Evolutionary Programming*, pages 345–353. Springer, Berlin, 1998.

[2] J. Ester. *Systemanalyse und mehrkriterielle Entscheidung*. VEB Verlag Technik, Berlin, 1987.

[3] W. T. Trotter. Partially ordered sets. In R. L. Graham, M. Grötschel, and L. Lovász, editors, *Handbook of Combinatorics, Vol. 1*, pages 433–480. Elsevier Science, Amsterdam, 1995.

[4] W. Klingenberg and P. Klein. *Lineare Algebra und analytische Geometrie, Band 1*. Bibliographisches Institut, Mannheim, 1971.

[5] F. A. Behringer. Zur Existenz maximaler Werte bei quasiordnungswertigen Abbildungen. *Mathematische Operationsforschung und Statistik, Series Optimization*, 10(5):193–205, 1979.

[6] G. Rudolph. Convergence of evolutionary algorithms in general search spaces. In *Proceedings of the Third IEEE Conference on Evolutionary Computation*, pages 50–54. IEEE Press, Piscataway (NJ), 1996.

[7] G. Rudolph. *Convergence Properties of Evolutionary Algorithms*. Kovač, Hamburg, 1997.

[8] M. Peschel and C. Riedel. Use of vector optimization in multiobjective decision making. In D. E. Bell, R. L. Keeney, and H. Raiffa, editors, *Conflicting Objectives in Decisions*, pages 97–121. Wiley, Chichester, 1977.

[9] G. Rudolph. On a multi–objective evolutionary algorithm and its convergence to the Pareto set. In *Proceedings of the 1998 IEEE International Conference on Evolutionary Computation*, pages 511–516. IEEE Press, Piscataway (NJ), 1998.

[10] T. Hanne. On the convergence of multiobjective evolutionary algorithms. *European Journal of Operational Research*, 1999 (to appear).