European Journal of Operational Research 000 (2014) 1–15

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

# Efficient optimization of many objectives by approximation-guided evolution

Markus Wagner[a,*], Karl Bringmann[b], Tobias Friedrich[c], Frank Neumann[a]

[a] University of Adelaide, Australia
[b] Max Planck Institute for Informatics, Saarbrücken, Germany
[c] Friedrich-Schiller-Universität Jena, Germany

ARTICLE INFO

*Article history:*
Received 3 November 2013
Accepted 18 November 2014
Available online xxxx

*Keywords:*
Multi-objective optimization
Approximation
Comparative study

ABSTRACT

Multi-objective optimization problems arise frequently in applications, but can often only be solved approximately by heuristic approaches. Evolutionary algorithms have been widely used to tackle multi-objective problems. These algorithms use different measures to ensure diversity in the objective space but are not guided by a formal notion of approximation. We present a framework for evolutionary multi-objective optimization that allows to work with a formal notion of approximation. This approximation-guided evolutionary algorithm (AGE) has a worst-case runtime linear in the number of objectives and works with an archive that is an approximation of the non-dominated objective vectors seen during the run of the algorithm. Our experimental results show that AGE finds competitive or better solutions not only regarding the achieved approximation, but also regarding the total hypervolume. For all considered test problems, even for many (i.e., more than ten) dimensions, AGE discovers a good approximation of the Pareto front. This is not the case for established algorithms such as NSGA-II, SPEA2, and SMS-EMOA. In this paper we compare AGE with two additional algorithms that use very fast hypervolume-approximations to guide their search. This significantly speeds up the runtime of the hypervolume-based algorithms, which now allows a comparison of the underlying selection schemes.



## 1. Introduction

Real-world optimization problems are usually very complex and hard to solve due to different circumstances such as constraints, complex function evaluations that can only be done by simulations, or multiple objectives. Most real-world optimization problems are characterized by multiple objectives. As these objectives are often in conflict with each other, the goal of solving a multi-objective optimization (MOO) problem is to find a (not too large) set of compromise solutions. The so-called Pareto front of a MOO problem consists of the function values representing the different trade-offs with respect to the given objective functions. The set of compromise solutions that is the outcome of a MOO run is an approximation of this Pareto front, and the idea of this posteriori approach is that afterwards the decision maker selects an efficient solution from this set. Multi-objective optimization is regarded to be more (or at least as) difficult as single-objective optimization due to the task of computing several solutions. From a computational complexity point of view even simple single-objective problems on weighted graphs like shortest paths or minimum spanning trees become NP-hard when they encounter at least two weight functions (Ehrgott, 2005). In addition, the size of the Pareto front is often exponential for discrete problems and even infinite for continuous ones.

Due to the hardness of almost all interesting multi-objective problems, different heuristic approaches have been used to tackle them. Among these methods, evolutionary algorithms are frequently used. They work at each time step with a set of solutions called population. The population of an evolutionary algorithm for a MOO is used to store desired trade-off solutions for the given problem.

As the size of the Pareto front is often very large, evolutionary algorithms and all other algorithms for MOO have to restrict themselves to a smaller set of solutions. This set of solutions should be a good approximation of the Pareto front. The main question is now how to define approximation. The literature (see e.g. Deb, 2001) on evolutionary multi-objective optimization (EMO) just states that the set of compromise solutions (i) should be close to the true Pareto front, (ii) should cover the complete Pareto front, and (iii) should be uniformly distributed. There are different evolutionary algorithms for multi-objective optimization such as NSGA-II (Deb, Pratap, Agrawal, & Meyarivan, 2002), SPEA2 (Zitzler, Laumanns, & Thiele, 2002), or IBEA (Zitzler & Künzli, 2004), which try to achieve these goals by preferring diverse sets of non-dominated solutions.

* Corresponding author. Tel.: +61 8 8313 5405; fax: +61 8 8313 4366.
*E-mail address:* markus.wagner@adelaide.edu.au, wagner@acrocon.com (M. Wagner).

http://dx.doi.org/10.1016/j.ejor.2014.11.032
0377-2217/

However, the above notion of approximation is not a formal definition. Having no formal definition of approximation makes it hard to evaluate and compare algorithms for MOO problems. Therefore, we think that it is necessary to use a formal definition of approximation in this context and evaluate algorithms with respect to this definition.

Different formal notions of approximation have been used to evaluate the quality of algorithms for multi-objective problems from a theoretical point of view. The most common ones are the multiplicative and additive approximations (see Cheng, Janiak, & Kovalyov, 1998; Daskalakis, Diakonikolas, & Yannakakis, 2010; Diakonikolas & Yannakakis, 2009; Papadimitriou & Yannakakis, 2000, 2001; Vassilvitskii & Yannakakis, 2005). Laumanns, Thiele, Deb, and Zitzler (2002) have incorporated this notion of approximation in an evolutionary algorithm for MOO. However, this algorithm is mainly of theoretical interest as the desired approximation is determined by a parameter of the algorithm and is not improved over time. Another approach related to a formal notion of approximation is the popular hypervolume indicator (Zitzler & Thiele, 1999) that measures the volume of the dominated portion of the objective space. Hypervolume-based algorithms such as MO-CMA-ES (Igel, Hansen, & Roth, 2007) or SMS-EMOA (Beume, Naujoks, & Emmerich, 2007) are well-established for solving MOO problems. They do not use a formal notion of approximation but it has recently been shown that the worst-case approximation obtained by optimal hypervolume distributions is asymptotically equivalent to the best worst-case approximation achievable by all sets of the same size (Bringmann & Friedrich, 2010b, 2010c). The major drawback of the hypervolume approach is that it cannot be computed in time polynomial in the number of objectives unless P = NP (Bringmann & Friedrich, 2010a). It is even NP-hard to determine which individual gives approximately the least contribution to the total hypervolume (Bringmann & Friedrich, 2012).

We introduce an efficient framework of an evolutionary algorithm for MOO that works with a formal notion of approximation and improves the approximation quality during its iterative process. The algorithm can be applied to a wide range of notions of approximation that are formally defined. As the algorithm does not have complete knowledge about the true Pareto front, it uses the best knowledge obtained so far during the optimization process.

The intuition for our algorithm is as follows. During the optimization process, the current best set of compromise solutions (usually called “population”) gets closer and closer to the Pareto front. Similarly, the set of all non-dominated points seen so far in the objective space (we call this “archive”) is getting closer to the Pareto front. Additionally, the archive is getting larger and larger and becoming an increasingly good approximation of the true Pareto front. Assuming that the archive approximates the Pareto front, we then measure the quality of the population by its approximation with respect to the archive. In our algorithm

- any set of feasible solutions constitutes an (potentially bad) approximation of the true Pareto front, and
- we optimize the approximation with respect to all solutions seen so far.

We introduce a basic approximation guided evolutionary algorithm which already performs very well for problems with many objectives. One drawback of the basic approach is that the archive size might grow tremendously during the run of the algorithm. In order to deal with this, we propose to work with an approximative archive which keeps at each time step only an $\epsilon$-approximation of all solutions seen so far. We do this by incorporating the $\epsilon$-dominance approach of Laumanns et al. (2002) into the algorithm. Furthermore, we introduce a powerful parent selection scheme which especially increases the performance of our algorithm for problems with just a few objectives by given the algorithm a stronger focus on the extreme points on the Pareto front.

We show on a set of well established benchmark problems that our approach is highly successful in obtaining high quality approximations according to the formal definition. Comparing our results to state of the art multi-objective algorithms such as NSGA-II, SPEA2, IBEA, and SMS-EMOA, we show that our algorithm typically gives better results, especially for high dimensional problems.

In our experimental study, we measure the quality of the results obtained not only in terms of the approximation quality but also with respect to the achieved hypervolume. Our experiments show that the examined hypervolume-based algorithms can sometimes achieve a larger hypervolume than our algorithm AGE, but AGE is the only one considered that finds a competitive hypervolume for all functions. Hence our algorithm not only performs better regarding our formal definition of approximation on problems with many objectives, but it is also competitive (or better, depending on the function) regarding the hypervolume.

This article is based on its previous conference publications. The based AGE algorithm has been introduced in Bringmann, Friedrich, Neumann, and Wagner (2011). The archive approximation has been presented in Wagner and Neumann (2013) and different parent selection schemes for AGE have been examined and discussed in Wagner and Friedrich (2013).

The outline of this paper is as follows. We introduce some basic definitions in Section 2. The main idea of approximation guided evolution and the basic AGE algorithm are presented in Section 3. In Section 6 we show how to speed up the approach by using an approximative archive and discuss different parent selection schemes in Section 5. We present our experimental results in Section 8 and finish with a summary and some concluding remarks.

## 2. Preliminaries

Multi-objective optimization deals with the optimization of several (often conflicting) objective functions. The different objective functions usually constitute a minimization or maximization problem on their own. Optimizing with respect to all given objective functions, there is usually no single optimal objective function vector, but a set of vectors representing the different trade-offs that are imposed by the objective functions.

Without loss of generality, we consider minimization problems with $d$ objective functions, where $d \geq 2$ holds. Each objective function $f_i : S \mapsto \mathbb{R}$, $1 \leq i \leq d$, maps from the considered search space $S$ into the real values. In order to simplify the presentation we only work with the dominance relation on the objective space and mention that this relation transfers to the corresponding elements of $S$.

For two points $x = (x_1, \ldots, x_d)$ and $y = (y_1, \ldots, y_d)$, with $x, y \in \mathbb{R}^d$ we define the following dominance relation:

$$x \preceq y \ :\Leftrightarrow \ x_i \leq y_i \text{ for all } 1 \leq i \leq d,$$
$$x \prec y \ :\Leftrightarrow \ x \preceq y \text{ and } x \neq y.$$

The typical notions of approximation used in theoretical computer science are multiplicative and additive approximation. We use the following definition

**Definition 1.** For finite sets $S, T \subset \mathbb{R}^d$, the additive approximation of $T$ with respect to $S$ is defined as

$$\alpha(S, T) := \max_{s \in S} \min_{t \in T} \max_{1 \leq i \leq d} (s_i - t_i).$$

In this paper, we only consider additive approximations. However, our approach can be easily adapted to multiplicative approximations. In this case, the term $s_i - t_i$ in Definition 1 has to be replaced by $s_i/t_i$.

## 3. Basic algorithm

Our aim is to minimize the additive approximation value of the population $P$ we output with respect to the archive $A$ of all points