A Convergence Criterion for Multiobjective Evolutionary Algorithms Based on Systematic Statistical Testing

Heike Trautmann[1], Uwe Ligges[2], Jörn Mehnen[3], and Mike Preuss[4]

[1] Department of Statistics, Technische Universität Dortmund
trautmann@statistik.uni-dortmund.de
[2] Department of Statistics, Technische Universität Dortmund
ligges@statistik.uni-dortmund.de
[3] Decision Engineering Centre, Cranfield University, UK
j.mehnen@cranfield.ac.uk
[4] Chair of Algorithm Engineering, Technische Universität Dortmund
mike.preuss@uni-dortmund.de

Abstract. A systematic approach for determining the generation number at which a specific Multi-Objective Evolutionary Algorithm (MOEA) has converged for a given optimization problem is introduced. Convergence is measured by the performance indicators Generational Distance, Spread and Hypervolume. The stochastic nature of the MOEA is taken into account by repeated runs per generation number which results in a highly robust procedure. For each generation number the MOEA is repeated a fixed number of times, and the Kolmogorow-Smirnov-Test is used in order to decide if a significant change in performance is gained in comparison to preceding generations. A comparison of different MOEAs on a problem with respect to necessary generation numbers becomes possible, and the understanding of the algorithm's behaviour is supported by analysing the development of the indicator values. The procedure is illustrated by means of standard test problems.

1 Introduction

Convergence properties for Multi-Objective Evolutionary Algorithms (MOEA) are an equally important issue as for single objective optimization. The question when to stop a stochastic search algorithm depends on practical as well as on technical decisions. Due to the fact that MOEAs are a fairly recent phenomenon there do not exist many mathematical convergence theories yet. Some of the current theories state back to single-objective theory (Deb 2004). Rudolph and Agapie (2000) and Rudolph (2001) proved that MOEAs with elitism and positive variation kernel can have the property of converging to the true Pareto front in finite number of function evaluations in finite search space problems. Further rigorous results are available for $t \to \infty$ (Hanne 1999). Laumanns et al. (2003, 2002) provided results on theoretical as well as empirical convergence properties for ϵ-MOEAs introducing additional spread properties. Van Veldhuizen and Lamont (1998a) derive results on sufficient conditions of convergence.

G. Rudolph et al. (Eds.): PPSN X, LNCS 5199, pp. 825–836, 2008.

In contrast to offline convergence analysis also online termination criteria for an MOEA exist, i.e. the MOEA is started for a single run, and it is stopped once a specific termination criterion has been met. In practice it is generally difficult to find a good termination criterion for MOEAs without sufficient a-priori knowledge about the optimization problem at hand. The most frequently used termination criterion is the maximum number of generations or execution time. An alternative is to measure the difference of improvements during a certain time interval. If the improvement is smaller than a certain threshold the termination criterion holds. The problem here lies in the determination of the threshold. Furthermore, this criterion may also be misleading in cases of functions with very small inclinations. Another criterion is to stop after a certain quality of a solution is reached. This leaves the problem of choosing a proper quality limit that allows for a finite termination of the algorithm. Rudenko and Schoenauer (2004) mention various online termination criteria for elite MOEAs. They discuss e.g. disappearance of all dominated individuals or deterioration of the number of newly produced non-dominated individuals. They propose a technique for determining stagnation using a stability measure of the crowding distance (Deb 2002). Deb and Jain (2002) investigate so-called running performance metrics for convergence and diversity of solutions to be monitored in the course of the algorithm.

Especially for problems with yet unknown characteristics a systematic analysis of the required run length of the MOEA becomes necessary as no sufficient a-priori-knowledge is available for choosing the desired solution quality. Termination criteria often are heuristical procedures with the disadvantage of being statistically unrobust as only a single run of the algorithm is taken into account.

In this paper a systematic offline convergence analysis of MOEA behaviour with respect to multiple performance indicators is suggested (Testing-based Runlength Detection (TRD)). Here Generational Distance (GD, Van Veldhuizen and Lamont 1998), Spread (Deb 2002) and Hypervolume (Zitzler and Thiele 1999) were chosen exemplary. For each generation number in a predefined (preliminary) interval the MOEA is applied m times resulting in m values of the performance indicators. This makes the procedure very robust as the stochastic nature of the MOEA is addressed.

Subsequently a Kolmogorow-Smirnov-Test (Sheskin 2000) is applied for each indicator in order to check if the distribution of the indicator at a specific generation number significantly differs from the distribution of the indicator values obtained at the five previous generations. The procedure stops in case the p-value of the test exceeds the significance level for three successive generations, and the indicator-specific optimal generation number is determined. The overall optimal generation number then comes out as the maximum of the indicator-specific optimal numbers. In case the preliminary upper generation limit has not been high enough it is redefined and the procedure is restarted beginning at the previous upper generation limit. Thus no computational recources are wasted.

By the proposed procedure it becomes possible to compare different MOEAs on an optimization problem in a systematic way with respect to the required generation numbers. Ideally the remaining MOEA-parameters should be adapted or tuned upfront for the given problem to allow highest possible algorithm performance. The method also ideally suits problems which have to be optimised repeatedly with the same or slightly