The Naive $\mathbb{MIDEA}$: A Baseline Multi–objective EA

Peter A.N. Bosman and Dirk Thierens

Institute of Information and Computing Sciences,
Utrecht University, The Netherlands
{Peter.Bosman, Dirk.Thierens}@cs.uu.nl

Abstract. Estimation of distribution algorithms have been shown to perform well on a wide variety of single–objective optimization problems. Here, we look at a simple - yet effective - extension of this paradigm for multi–objective optimization, called the naive $\mathbb{MIDEA}$. The probabilistic model in this specific algorithm is a mixture distribution, and each component in the mixture is a univariate factorization. Mixture distributions allow for wide–spread exploration of the Pareto front thus aiding the important preservation of diversity in multi–objective optimization. Due to its simplicity, speed, and effectiveness the naive $\mathbb{MIDEA}$ can well serve as a baseline algorithm for multi–objective evolutionary algorithms.

1 Introduction

Estimation of distribution algorithms (EDAs) are a class of evolutionary algorithms that build at each generation a probabilistic model from the selected solutions. EDAs are mainly characterized by the kind of probabilistic model they learn. Somewhat surprisingly, it has become clear that even the use of a simple univariate factorization as probabilistic model often leads to good performance for discrete, single objective optimization problems. Considering their ease of implementation, univariate EDAs have become a baseline algorithm that can be used to generate reasonable good solutions quickly, thus setting a performance level that more elaborated algorithms need to surpass to justify their use. The goal of this paper is to show that a similar baseline algorithm can be constructed for multi–objective optimization problems. The resulting algorithm - called the naive $\mathbb{MIDEA}$ - is a simple, fast, and efficient multi–objective evolutionary algorithm (MOEA) based on the concept of Pareto dominance, and can thus be used as a baseline algorithm for more elaborate MOEAs.

The remainder of this paper is organized as follows. In Section 2 we specify the naive $\mathbb{MIDEA}$ algorithm as an instance of the more general $\mathbb{MIDEA}$ (mixture–based, multi–objective iterated density–estimation evolutionary algorithm) framework. In Section 3 we test the performance of the algorithm on four multi–objective, combinatorial optimization problems, compare the results with two state–of–the–art MOEAs, and discuss our findings. We present our conclusions in Section 4.

C. A. Coello Coello et al. (Eds.): EMO 2005, LNCS 3410, pp. 428–442, 2005.

2 The Naive MIDEA

2.1 Mixture Probability Distributions

A mixture probability distribution is a weighted sum of k probability distributions. Each probability distribution in the mixture distribution is called a mixture component. Let $\boldsymbol{\mathcal{Z}} = (Z_0, Z_1, \ldots, Z_{l-1})$ be a vector of random variables Z_i associated with the i–th problem variable. A mixture probability distribution is then defined as:

$$P^{mixture}(\boldsymbol{\mathcal{Z}}) = \sum_{i=0}^{k-1} \beta_i P^i(\boldsymbol{\mathcal{Z}}) \tag{1}$$

where $\beta_i > 0$, $i \in \{0, 1, \ldots, k-1\}$, and $\sum_{i=0}^{k-1} \beta_i = 1$. The β_i with which the mixture components are weighted in the sum are called mixing coefficients.

The general advantage of mixture probability distributions is that a larger class of dependency relations between the random variables can be expressed than when using non–mixture probability distributions since a mixture probability distribution makes a combination of multiple probability distributions. In many cases, simple probability distributions can be used as mixture components to get accurate descriptions of the data in different parts of the sample space. By using mixture probability distributions, a powerful, yet computationally tractable type of probability distribution can be used within EDAs, that provides for processing complicated interactions between a problem's variables.

For multi–objective optimization, mixture distributions have an additional advantage that renders them particularly useful. The specific advantage is geometrical in nature. If we, for instance, cluster the solutions as observed in the objective space and then estimate a simpler probability distribution in each cluster, the probability distributions in these clusters can portray specific information about the different regions along the Pareto optimal front that we are ultimately interested in. Drawing new solutions from the resulting mixture probability distribution gives solutions that are more likely to be well spread along the front as each mixture component delivers a subset of new solutions. The use of such a mixture distribution thus results in a parallel exploration along the current Pareto front. This parallel exploration may very well provide a better spread of new solutions along the Pareto front than when a single non–mixture distribution is used to capture information about the complete Pareto front.

To complete the construction of the mixture distribution we also need to determine the mixing coefficients β_i. This can be done in various ways. Here we set β_i proportional to the size of the i–th cluster with respect to the sum of the sizes of all clusters.

2.2 Selection Operator

In multi-objective optimization we want the solutions to be as close to the Pareto optimal front as possible, and we want a good diverse representation of the Pareto optimal front. In a practical application, we have no indication of how close we are to the Pareto optimal front. To ensure selection pressure toward the