Operations Research Letters 39 (2011) 150–154

Contents lists available at ScienceDirect

Operations Research Letters

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

MB-GNG: Addressing drawbacks in multi-objective optimization estimation of distribution algorithms

Luis Martí [a,*], Jesús García [a], Antonio Berlanga [a], Carlos A. Coello Coello [b], José M. Molina [a]

[a] Group of Applied Artificial Intelligence, Department of Informatics, Universidad Carlos III de Madrid. Av. de la Universidad Carlos III, 22. Colmenarejo 28270 Madrid, Spain

[b] Department of Computer Science, CINVESTAV-IPN, Av. IPN No. 2508, Col. San Pedro Zacatenco México, D.F. 07360, Mexico

ARTICLE INFO

Article history:
Received 12 July 2010
Accepted 15 January 2011
Available online 1 February 2011

Keywords:
Multi-objective optimization
Estimation of distribution algorithm
Model building
Growing neural gas

ABSTRACT

We examine the model-building issue related to multi-objective estimation of distribution algorithms (MOEDAs) and show that some of their, as yet overlooked, characteristics render most current MOEDAs unviable when addressing optimization problems with many objectives. We propose a novel model-building growing neural gas (MB-GNG) network that is specially devised for properly dealing with that issue and therefore yields a better performance. Experiments are conducted in order to show from an empirical point of view the advantages of the new algorithm.

1. Introduction

Most human endeavors involve the creation of artifacts with properties that must be tuned to be as efficient as possible. This fact has prompted the creation of a number of interrelated research areas like optimization, mathematical programming, operational research and decision-making. Although these areas share some of their goals, each of them differs from the others in the approaches put forward by their respective communities and the characteristics of the problems they deal with.

Many real-world optimization problems involve more than one goal to be optimized. This type of problems is known as *multi-objective optimization problems* (MOPs). A MOP can be expressed as the problem in which a set of *objective functions* $f_1(\boldsymbol{x}), \ldots, f_M(\boldsymbol{x})$ should be jointly optimized;

$$\min \boldsymbol{F}(\boldsymbol{x}) = \langle f_1(\boldsymbol{x}), \ldots, f_M(\boldsymbol{x}) \rangle; \quad \boldsymbol{x} \in \mathcal{S}; \tag{1}$$

where $\mathcal{S} \subseteq \mathbb{R}^n$ is known as the *feasible set* and could be expressed as a set of restrictions over the decision set, $\mathbb{R}^n$. The image set of $\mathcal{S}$ produced by function vector $\boldsymbol{F}(\cdot)$, $\mathcal{O} \subseteq \mathbb{R}^M$, is called *feasible objective set* or *criterion set*.

The solution to this type of problem is a set of trade-off points. The optimality of a solution can be expressed in terms of the Pareto dominance relation.

Definition 1 (*Pareto Dominance Relation*)**.** For the optimization problem specified in (1) and having $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{S}$, $\boldsymbol{x}$ is said to dominate $\boldsymbol{y}$ (expressed as $\boldsymbol{x} \prec \boldsymbol{y}$) iff $\forall f_j$, $f_j(\boldsymbol{x}) \leq f_j(\boldsymbol{y})$ and $\exists f_i$ such that $f_i(\boldsymbol{x}) < f_i(\boldsymbol{y})$.

Definition 2 (*Non-Dominated Subset*)**.** In problem (1) and having the set $\mathcal{A} \subseteq \mathcal{S}$. $\hat{\mathcal{A}}$, the *non-dominated subset* of $\mathcal{A}$, is defined as

$$\hat{\mathcal{A}} = \left\{ \boldsymbol{x} \in \mathcal{A} \mid \nexists \boldsymbol{x}' \in \mathcal{A} : \boldsymbol{x}' \prec \boldsymbol{x} \right\}.$$

The solution of (1) is $\hat{\mathcal{S}}$, the non-dominated subset of $\mathcal{S}$. $\hat{\mathcal{S}}$ is known as the *efficient set* or *Pareto-optimal set* [4]. If problem (1) has certain characteristics, e.g., linearity or convexity of the objective functions or convexity of $\mathcal{S}$, the efficient set can be determined by mathematical programming approaches [4]. However, in the general case, finding the solution of (1) is an *NP*-complete problem [2]. In this case, heuristic or metaheuristic methods can be applied in order to have solutions of practical value at an admissible computational cost.

A broad range of heuristic and metaheuristic approaches has been used to address MOPs [4]. Of these, multi-objective evolutionary algorithms (MOEAs) [5] have been found to be a competent approach in a wide variety of application domains. Their main advantages are ease of use, inherent parallel search and lower susceptibility to the shape or continuity of the image of the efficient set, compared with traditional mathematical programming techniques for multi-objective optimization [4].

There is a class of MOPs that are particularly appealing because of their inherent complexity: the so-called *many-objective problems*. These are problems with a relatively large number of objectives (normally, four or more). Although somewhat counter-intuitive and hard to visualize for a human decision maker, these problems are not uncommon in real-life engineering practice. For example, [14] details some relevant real problems of this type.

* Corresponding author.
E-mail addresses: lmarti@inf.uc3m.es (L. Martí), jgherrer@inf.uc3m.es (J. García), aberlan@ia.uc3m.es (A. Berlanga), ccoello@cs.cinvestav.mx (C.A. Coello Coello), molina@ia.uc3m.es (J.M. Molina).

0167-6377/$ – see front matter
doi:10.1016/j.orl.2011.01.002

The scalability issues of traditional MOEAs in these problems have triggered a sizable amount of research, aiming to provide alternative approaches that can properly handle many-objective problems and perform reasonably well.

Estimation of distribution algorithms (EDAs) are one such approaches [11]. EDAs have been hailed as a paradigm shift in evolutionary computation. They build a model of the population instead of applying evolutionary operators. This model is then used to synthesize new individuals. EDAs have been extended to the multi-objective optimization problem domain as multi-objective EDAs (MOEDAs).

Although MOEDAs have yielded some encouraging results, their introduction has not lived up to *a priori* expectations. This can be attributed to a number of different causes. We have recognized three of them, in particular, those derived from the incorrect treatment of population outliers; the loss of population diversity, and that too much computational effort is being spent on finding an optimal population model.

A number of works have dealt with the issues listed above, particularly with loss of diversity. Nevertheless, in our opinion, the community has failed to acknowledge that the underlying cause for all those problems could, perhaps, be traced back to the algorithms used for model building in EDAs.

In this paper we examine the model-building issue of current MOEDAs and show that some of its characteristics, which have been disregarded so far, render most current approaches unsuitable for tackling MOPs. We then propose a novel model-building algorithm, based on the growing neural gas (GNG) network. This model-building GNG (MB-GNG) is the main contribution of this paper. It has been devised with this particular problem in mind, and therefore addresses the problems of current approaches.

The remainder of this paper is organized as follows. Section 2 serves as a brief introduction to MOEDAs and the issues present in current model-building algorithms. After this, MB-GNG is described in Section 4. Then, in Section 5, a comparative study is carried out in order to establish from an experimental point of view the improvements introduced by MB-GNG with respect to similar algorithms. Finally, some conclusive remarks are put forward.

2. Multi-objective estimation of distribution algorithms

Estimation of distribution algorithms (EDAs) are population-based optimization algorithms that rely on machine learning methods. The introduction of machine learning implies that these new algorithms lose the straightforward biological inspiration of their predecessors. Nonetheless, they gain the capacity of scalably solving many challenging problems, in some cases significantly outperforming standard EAs and other optimization techniques.

Most multi-objective EDAs (MOEDAs) consist of a modification of existing EDAs whose fitness assignment function is substituted by one taken from an existing MOEA.

In general terms, MOEDAs follow a common algorithmic scheme. At a given iteration t, a MOEDA has a population $\mathcal{P}_t$ of individuals, each one representing a point in the search space. In every iteration, $\mathcal{P}_t$ elements are ranked according to a given fitness assignment function. A subset $\mathcal{M}_t$, with the best elements of $\mathcal{P}_t$ is computed. The model-building algorithm relies on $\mathcal{M}_t$ to create a model of the best part of the population. This model is sampled in order to create new elements which are combined with $\mathcal{P}_t$ to create the population to be used in the next iteration, $\mathcal{P}_{t+1}$.

A given stopping criterion determines when the optimization process should be interrupted. When this happens, $\hat{\mathcal{P}}$, the non-dominated subset of $\mathcal{P}_t$, is returned as the solution.

Although there are different approaches for determining $\mathcal{M}_t$ and $\mathcal{P}_{t+1}$, MOEDAs are better characterized by their two main components, the fitness assignment function and the model-building algorithm.

Fitness functions have been mostly taken from MOEAs. It should be noted that the Pareto dominance-based approach proposed by the NSGA-II algorithm is, by far, the most popular in the current literature.

The model-building algorithm is the kernel of an EDA. There are two main types of methods for addressing this problem: those based on graphical models and those based on mixture distributions.

2.1. Graphical model MOEDAs

Most EDAs based on graphical models rely on Bayesian networks. From these, the Bayesian optimization algorithm (BOA) [11] is the specific approach that has been extrapolated to the multi-objective domain. The exhaustive synthesis of a Bayesian network from the algorithm's population is an NP-hard problem [7]. Therefore, these EDAs must employ heuristic alternatives for building their networks while keeping the computational cost under reasonable margins.

BOA-based MOEDAs combine the Bayesian model-building scheme with an already existing Pareto-based fitness assignment. This is the case of the multi-objective BOA (mBOA) that exploits the fitness assignment used in NSGA-II. Another algorithm based on hierarchical BOA (hBOA), called mhBOA, also uses the same form of fitness assignment but introduces clustering in the feasible objective set. A similar idea is proposed by combining the mixed BOA (mBOA) with SPEA2's selection scheme [5] to form the multi-objective mBOA (mmBOA). The multi-objective real BOA (MrBOA) [1] also extends a preexisting EDA, namely, the real BOA (rBOA). MrBOA combines the fitness assignment of NSGA-II with rBOA.

2.2. Mixture distribution MOEDAs

Another approach to modeling the subset with the best population elements is to apply a distribution mixture approach. Bosman and Thierens [3] proposed several variants of their multi-objective mixture-based iterated density estimation algorithm (MIDEA). They are based on their IDEA framework. They also introduced a novel Pareto-based and diversity-preserving fitness assignment function. The model construction is inherited from the single-objective version. The proposed MIDEAs considered several types of probabilistic models for both discrete and continuous problems.

MIDEAs do not provide a specific mechanism to ensure equal coverage of the Pareto-optimal front if the number of representatives in some parts of the front is much larger than the number of representatives in some other parts. The clustering algorithms applied for this task include the randomized leader algorithm, the k-means algorithm and the expectation maximization (EM) algorithm [15].

2.3. Other MOEDA approaches

There are some other approaches for model building. For example, the regularity model-based multi-objective estimation of distribution algorithm (RM-MEDA) [16] is based on the regularity property derived from the Karush–Kuhn–Tucker condition. Covariance matrix adaptation evolution strategies (CMA-ES) [9] have been also used in the multi-objective context. CMA-ES consists of a method for updating the covariance matrix of the multivariate normal mutation distribution used in an evolution strategy. They can be viewed as EDAs, as new individuals are sampled according to the mutation distribution.