A software framework based on a conceptual unified model for evolutionary multiobjective optimization: ParadisEO-MOEO

Arnaud Liefooghe, Laetitia Jourdan, El-Ghazali Talbi

To cite this version:


HAL Id: hal-00522612
https://hal.archives-ouvertes.fr/hal-00522612
Submitted on 1 Oct 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Decision Support

A software framework based on a conceptual unified model for evolutionary multiobjective optimization: ParadisEO-MOEO

Arnaud Liefooghe a,b,*, Laetitia Jourdan a, El-Ghazali Talbi a,c

aUniversité Lille 1, Laboratoire d’Informatique Fondamentale de Lille, UMR CNRS 8022, INRIA Lille-Nord Europe, Parc Scientifique de la Haute Borne, 40 Avenue Halley, 59650 Villeneuve d’Ascq, France
bUniversity of Coimbra, CESUC, Department of Informatics Engineering, Portugal
cKing Saud University, Riyadh, Saudi Arabia

1. Introduction

Evolutionary multiobjective optimization (EMO) is one of the most challenging areas in the field of multicriteria decision making. Generally speaking, a multiobjective optimization problem (MOP) can be defined by a vector function \( f \) of \( n \geq 2 \) objective functions \( f_1, f_2, \ldots, f_n \), and a set \( X \) of feasible solutions in the decision space. Let \( Z \) be the set of feasible points in the objective space. Without loss of generality, we assume that \( Z \subseteq \mathbb{R}^n \) and that all \( n \) objective functions are to be minimized. To each decision vector \( x \in X \) is assigned an objective vector \( z = f(x) \). A dominance relation is then usually assumed so that a partial order is induced over the set \( X \) of feasible solutions in the decision space. With this order, we can define a set \( P \subseteq X \) of non-dominated points in \( X \). This set is called the Pareto optimal set.

Keywords:
- Multiple objective programming
- Evolutionary algorithms
- Conceptual unified model
- Algorithm design and implementation
- Software framework

abstract

This paper presents a general-purpose software framework dedicated to the design and the implementation of evolutionary multiobjective optimization techniques: ParadisEO-MOEO. A concise overview of evolutionary algorithms for multiobjective optimization is given. A substantial number of methods has been proposed so far, and an attempt of conceptually unifying existing approaches is presented here. Based on a fine-grained decomposition and following the main issues of fitness assignment, diversity preservation and elitism, a conceptual model is proposed and is validated by a number of state-of-the-art algorithms as simple variants of the same structure. This model is then incorporated into the ParadisEO-MOEO software framework. This framework has proven its validity and high flexibility by enabling the resolution of many academic, real-world and hard multiobjective optimization problems.

© 2010 Elsevier B.V. All rights reserved.

* Corresponding author at: Université Lille 1, Laboratoire d’Informatique Fondamentale de Lille, UMR CNRS 8022, INRIA Lille-Nord Europe, Parc Scientifique de la Haute Borne, 40 Avenue Halley, 59650 Villeneuve d’Ascq, France. Tel.: +33 3 59 57 79 30; fax: +33 3 59 57 78 50.

E-mail addresses: arnaud.liefooghe@lifl.fr (A. Liefooghe), laetitia.jourdan@lifl.fr (L. Jourdan), talbi@lifl.fr (E.-G. Talbi).

As pointed out by different authors (see, e.g., Coello Coello et al., 2007; Zitzler et al., 2004), approximating an efficient set is itself a...
bi-objective problem. Indeed, the approximation to be found must have both good convergence and distribution properties, as its mapping in the objective space has to be (i) close to and (ii) well-spread over the (generally unknown) optimal Pareto front, or a subpart of it. As a consequence, the main difference between the concept of a single-objective and of a multiobjective search method deals with these two goals. Over the last two decades, major advances, from both algorithmic and theoretical aspect, have been made in the EMO field. And a large number of algorithms have been proposed. Among the existing approaches, one may cite VEGA (Schaffer, 1985), MOGA (Fonseca and Fleming, 1993), NSGA (Srinivas and Deb, 1994), NSGA-II (Deb et al., 2002), NPGA (Horn et al., 1994), SPEA (Zitzler and Thiele, 1999), SPEA2 (Zitzler et al., 2001) or PESA (Corne et al., 2000). Note that another topic to mention while dealing with EMO relates to performance assessment. Various quality indicators have been proposed in the literature for evaluating the performance of multiobjective search methods. The reader is referred to Zitzler et al. (2003) for a review.

Zitzler et al. (2004) notice that initial EMO approaches were mainly focused on moving toward the Pareto front (Schaffer, 1985; Fourman, 1985). Afterwards, diversity preservation mechanisms quickly emerged (Fonseca and Fleming, 1993; Srinivas and Deb, 1994; Horn et al., 1994). Then, at the end of the 90s, the concept of elitism, related to the preservation and use of non-dominated solutions, became very popular and is now employed in most recent EMO methods (Zitzler and Thiele, 1999; Zitzler et al., 2001; Knowles and Corne, 2000). Specific issues of fitness assignment, diversity preservation and elitism are commonly approved in the community and are also presented under different names in, for instance (Coello Coello et al., 2007; Zitzler et al., 2004). Based on these three main notions, several attempts have been made in the past for unifying EMO algorithms under ParadisEO-MOEO. Finally, the last section concludes the paper.

2. A conceptual unified model for evolutionary multiobjective optimization

An evolutionary algorithm (EA) (Eiben and Smith, 2003) is a search method that belongs to the class of metaheuristics (Talbi, 2009), and where a population of solutions is iteratively improved by means of some stochastic operators. Starting from an initial population, each individual is evaluated in the objective space and a selection scheme is performed to build a so-called parent population. An offspring population is then created by applying variation operators. Next, a replacement strategy determines which individuals will survive. This search process is iterated until a given stopping condition is satisfied.

As noticed earlier in the paper, in the frame of EMO, the main expansions deal with the issues of fitness assignment, diversity preservation and elitism. Indeed, contrary to single-objective optimization where the fitness value of a solution corresponds to its single-objective value in most cases, a multiobjective fitness assignment scheme is here required to assess the individuals’ performance, as the mapping of a solution in the objective space is now multidimensional. Moreover, trying to approximate the efficient set is not only a question of convergence. The approximation to be found also has to be well-spread over the objective space, so that a diversity preservation mechanism is usually required. This fitness and diversity information are required to discriminate individuals at the selection and the replacement steps of the EA. Next, the main purpose of elitism is to avoid the loss of best-found non-dominated solutions during the stochastic search process. These solutions are frequently incorporated into a secondary population, the so-called archive. The update of the archive contents possibly appear at each EA iteration.

As a consequence, whatever the MOP to be solved, the common concepts for the design of an EMO algorithm are the following ones:

1. Design a representation.
2. Design a population initialization strategy.
3. Design a way of evaluating a solution.
4. Design suitable variation operators.
5. Decide a fitness assignment strategy.
6. Decide a diversity preservation strategy.
7. Decide a selection strategy.
8. Decide a replacement strategy.
9. Decide an archive management strategy.
10. Decide a continuation strategy.
When dealing with any kind of metaheuristics, one may distinguish problem-specific and generic parts. Indeed, the former four common concepts presented above strongly depend on the MOP at hand, while the six latter ones can be considered as problem independent, even if some problem-dependent strategies can also be designed in some particular cases. Note that the concepts of representation and evaluation are shared by any metaheuristic, the concepts of population initialization and stopping criterion are shared by any population-based metaheuristic, the concepts of variation operators, selection and replacement are shared by any EA, whereas the concepts of fitness, diversity and archiving are specific to EMO.

2.1. Elements description

This section provides a description of the elements involved in our conceptual unified model. EMO-related issues are detailed in more depth.

2.1.1. Representation

Solution representation is the starting point for anyone who plans to design any kind of metaheuristic. An MOP solution needs to be represented both in the decision space and in the objective space. While the representation in the objective space can be seen as problem independent, the representation in the decision space must be relevant to the tackled problem. Successful applications of metaheuristics strongly require a proper solution representation. Various encodings may be used such as binary variables, real-coded vectors, permutations, discrete vectors, and more complex representations. The choice of a representation will considerably influence the way solutions will be initialized and evaluated in the objective space, and the way variation operators will be applied.

2.1.2. Initialization

Whatever the algorithmic solution to be designed, a way to initialize a solution (or a population of solutions) is expected. While dealing with any population-based metaheuristic, one has to keep in mind that the initial population must be well diversified in order to prevent a premature convergence. This remark is even more true for MOPs where the goal is to find a well-converged and well-spread approximation. The way to initialize a solution is closely related to the problem under consideration and to the representation at hand. In most approaches, the initial population is generated randomly or according to a given diversity function.

2.1.3. Evaluation

The problem at hand is to optimize a set of objective functions simultaneously over a given search space. Then, each time a new solution integrates the population, its objective vector must be set, i.e., the value corresponding to each objective function must be evaluated.

2.1.4. Variation

The purpose of variation operators is to modify the representation of solutions in order to move them in the search space. Generally speaking, while dealing with EAs, these problem-dependent operators are stochastic. Mutation operators are unary operators acting on a single solution whereas recombination (or crossover) operators are mostly binary, and sometimes n-ary.

2.1.5. Fitness assignment

In the single-objective case, the fitness value of a solution is most often its unidimensional objective value. While dealing with MOPs, fitness assignment aims to guide the search toward Pareto optimal solutions for a better convergence. Extending the works of Zitzler et al. (2004) and Coello Coello et al. (2007), we propose to classify existing fitness assignment schemes into four different families:

- **Scalar approaches**, where the MOP is reduced to a single-objective optimization problem. A popular example consists of combining the $n$ objective functions into a single one by means of a weighted-sum aggregation. Other examples are $c$-constraint or achievement function-based methods (see Miettinen, 1999).

- **Criterion-based approaches**, where each objective function is treated separately. For instance, in VEGA (vector evaluated genetic algorithm) (Schaffer, 1985), a parallel selection is performed where solutions are discerned according to their values on a single-objective function, independently to the others. In lexicographic methods (Fourman, 1985), a hierarchical order is defined beforehand between objective functions.

- **Dominance-based approaches**, where a dominance relation is used to classify solutions. For instance, dominance-rank techniques compute the number of population items that dominate a given solution (Fonseca and Fleming, 1993). Such a strategy takes part in, e.g., Fonseca and Fleming MOGA (multiobjective genetic algorithm) (Fonseca and Fleming, 1993). In dominance-count techniques, the fitness value of a solution corresponds to the number of individuals that are dominated by these solutions (Zitzler and Thiele, 1999). Finally, dominance-depth strategies consist of classifying a set of solutions into different classes (or fronts) (Goldberg, 1989). Hence, a solution that belongs to a class does not dominate another one from the same class; so that individuals from the first front all belong to the best non-dominated set, individuals from the second front all belong to the second best non-dominated set, and so on. The latter approach is used in both NSGA (non-dominated sorting genetic algorithm) (Srinivas and Deb, 1994) and NSGA-II (Deb et al., 2002). In addition, note that several schemes can also be combined, as in the case of, for example, Zitzler and Thiele (1999).

In the frame of dominance-based approaches, the most commonly used dominance relation is based on Pareto-dominance as given in Section 1. But some recent techniques are based on other dominance operators such as $\epsilon$-dominance in Deb et al. (2005a) and $g$-dominance in Molina et al. (2009).

- **Indicator-based approaches**, where the fitness values are computed by comparing individuals on the basis of a quality indicator $I$. The chosen indicator represents the overall goal of the search process. Generally speaking, no particular diversity preservation mechanism is in usual necessary, with regards to the indicator being used. Examples of indicator-based EAs are IBEA (indicator-based EA) proposed by Zitzler and Künzli (2004) or SMS-EMOA (S-metric selection EMO algorithm) of Beume et al. (2007).

2.1.6. Diversity assignment

As noticed in the previous section, aiming at approximating the efficient set is not only an issue related to convergence. The final approximation also has to be well-spread over the objective space. However, classical dominance-based fitness assignment schemes often tend to produce premature convergence by privileging non-dominated solutions, what does not guarantee a uniformly sampled output set. In order to prevent that issue, a diversity preservation mechanism, based on a given distance measure, is usually integrated into the algorithm to uniformly distribute the population over the trade-off surface. In the frame of EMO, a common distance measure is based on the euclidean distance between objective vectors. But, this measure can also be defined in the decision space or can even combined both spaces. Nevertheless, a distance metric partly or fully defined in the parameter space strongly relies on the tackled problem.

---

Popular examples of EMO diversity assignment techniques are sharing or crowding. The notion of sharing (or fitness sharing) has initially been suggested by Goldberg and Richardson (1987) to preserve diversity among the solutions of an EA population. It has first been employed by Fonseca and Fleming (1993) in the frame of EMO. This kernel method consists of estimating the distribution density of a solution using a so-called sharing function that is related to the sum of distances to its neighborhood solutions. A sharing distance parameter specifies the similarity threshold, i.e., the size of niches. Another diversity assignment scheme is the concept of crowding, first suggested by Holland (1975) and used by De Jong (1975) to prevent genetic drift. It is employed by Deb et al. (2002) in the frame of NSGA-II. Contrary to sharing, this scheme allows to maintain diversity without specifying any parameter. It consists in estimating the density of solutions surrounding a particular point of the objective space.

2.1.7. Selection

The selection step is one of the main search operators of EAs. It consists of choosing some solutions that will be used to generate the offspring population. In general, the better an individual, the higher its chance of being selected. Common strategies are deterministic or stochastic tournament, roulette-wheel selection, random selection, etc. An existing EMO-specific elitist scheme consists of including solutions from the archive in the selection process, so that non-dominated solutions also contribute to the evolution engine. Such an approach has successfully been applied in various elitist EMO algorithms including SPEA (Zitzler and Thiele, 1999), SPEA2 (Zitzler et al., 2001) and PESA (Corne et al., 2000). In addition, in order to prohibit the crossover of dissimilar parents, mating restriction (Goldberg, 1989) can also be mentioned as a candidate strategy to be integrated into EMO algorithms.

2.1.8. Replacement

Selection pressure is also affected at the replacement step where survivors are selected from both the current and the offspring populations. In generational replacement, the offspring population systematically replaces the parent one. An elitist strategy consists of preserving the $N$ best solutions from both populations ($N$ stands for the appropriate population size), either by means of a one-shot deletion or by deleting the worst solution iteratively until the required population size is reached, so that fitness and diversity information of remaining solutions can be updated each time there is a deletion.

2.1.9. Elitism

Another essential issue about MOP solving is the notion of elitism. It mainly consists of maintaining an external set, the so-called archive, that allows to store either all or a subset of non-dominated solutions found during the search process. This secondary population mainly aims at preventing the loss of these solutions during the stochastic optimization process. The update of the archive contents with new potential non-dominated solutions is mostly based on the Pareto-dominance relation. But, in the literature, other dominance criterion can be found. Examples are weak-dominance, strict-dominance, $\epsilon$-dominance (Helbig and Pateva, 1994; Laumanns et al., 2002), and so on. When dealing about archiving, one may distinguish four different techniques depending on the problem properties, the designed algorithm and the number of desired solutions: (i) no archive, (ii) an unbounded archive, (iii) a bounded archive or (iv) a fixed-size archive. Firstly, if the current approximation is maintained by, or contained in the main population itself, there can be no archive at all. On the other hand, if an archive is maintained, it usually comprises the current non-dominated set approximation, as dominated solutions are removed. Then, an unbounded archive can be used in order to save the whole set of non-dominated solutions found until the beginning of the search process. However, as some continuous optimization problems may contain an infinite number of non-dominated solutions, it is simply not possible to save them all. Therefore, additional operations must be used to reduce the number of stored solutions. Then, a common strategy is to bound the size of the archive according to some fitness and/or diversity assignment scheme(s). Finally, another archiving technique consists of a fixed size storage capacity, where a bounding mechanism is used when there is too much non-dominated solutions, and some dominated solutions are integrated in the archive if the non-dominated set is too small, what is done for instance within SPEA2 (Zitzler et al., 2001). Usually, an archive is used as an external storage only. However, archive members can also be integrated during the selection phase of an EMO algorithm, as proposed by Zitzler and Thiele (1999), see Section 2.1.7.

2.1.10. Stopping criteria

Since an iterative method computes successive approximations, a practical test is required to determine when the process must stop. Popular examples are a given number of iterations, a given number of evaluations, a given run time, etc.

2.2. State-of-the-art EMO algorithms as instances of the conceptual unified model

By means of the conceptual model presented in this paper, we claim that a large number of state-of-the-art EMO algorithms proposed in the last two decades are based on variations of the problem-independent elements presented above. In Table 1, three EMO approaches, namely NSGA-II (Deb et al., 2002), SPEA2 (Zitzler et al., 2001) and IBEA (Zitzler and Künzli, 2004), are regarded as simple instances of the model proposed in this paper. For obvious reasons, only problem-independent issues are presented. NSGA-II and SPEA2 are two of the most frequently encountered EMO algorithms of the literature, either for tackling an original MOP or to serve as references for comparison. Regarding IBEA, it is a good illustration of the new EMO trend dealing with indicator-based search that started to become popular in recent years. We can see in the table that these three state-of-the-art algorithms perfectly fit into our model for EMO, what strongly validates the proposed approach. But other examples can be found in the literature. For instance, the only part that differs from NSGA (Srinivas and Deb, 1994) to NSGA-II is the diversity preservation strategy that is based on sharing in NSGA and on crowding in NSGA-II. Another example is the c-MOEA proposed by Deb et al. (2005a). This algorithm is a

<table>
<thead>
<tr>
<th>Elements</th>
<th>NSGA-II</th>
<th>SPEA2</th>
<th>IBEA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fitness assignment</td>
<td>Dominance-depth</td>
<td>Dominance-count and rank</td>
<td>Binary quality indicator</td>
</tr>
<tr>
<td>Diversity assignment</td>
<td>Crowding distance</td>
<td>kth nearest neighbor</td>
<td>None</td>
</tr>
<tr>
<td>Selection</td>
<td>Binary tournament</td>
<td>Elitist binary tournament</td>
<td>Binary tournament</td>
</tr>
<tr>
<td>Replacement</td>
<td>One-shot elitist replacement</td>
<td>Generational replacement</td>
<td>Iterative elitist replacement</td>
</tr>
<tr>
<td>Archiving</td>
<td>None</td>
<td>Fixed-size archive</td>
<td>Number of generations</td>
</tr>
<tr>
<td>Stopping criteria</td>
<td>Number of generations</td>
<td>Number of generations</td>
<td>Number of generations</td>
</tr>
</tbody>
</table>

Table 1: State-of-the-art EMO methods as instances of the proposed unified model.

modified version of NSGA-II where the Pareto-dominance relation used for fitness assignment has been replaced by the ϵ-dominance relation. Similarly, the g-dominance relation proposed by Molina et al. (2009) is experimented by the authors on an NSGA-II-like EMO technique where the dominance relation has been modified in order to take the DM preferences into account by means of a reference point.

3. Design and implementation of evolutionary multiobjective optimization algorithms under ParadisEO-MOEO

In this section, we provide a general presentation of ParadisEO, a software framework dedicated to the design of metaheuristics, and a detailed description of the ParadisEO module specifically dedicated to EMO, namely ParadisEO-MOEO. Historically, ParadisEO was especially dedicated to parallel and distributed metaheuristics and was the result of the Ph.D. work of Sébastien Cahon, supervised by Nouredine Melab and El-Ghazali Talbi (Cahon et al., 2004). The initial version already contained a few number of EMO-related features, mainly with regard to archiving. This work has been partially extended and presented in Liefooghe et al. (2007b). But since then, the ParadisEO-MOEO module has been completely redesigned in order to confer an even more fine-grained decomposition in accordance with the conceptual model presented above.

3.1. Motivations

In practice, there exists a large diversity of optimization problems to be solved, engendering wide possibilities in terms of models to handle in the frame of a metaheuristic solution method. Moreover, a growing number of general-purpose search methods are proposed in the literature, with evolving complex mechanisms. From a practitioner point of view, there is a popular demand to provide a set of ready-to-use metaheuristic implementations, allowing a minimum programming effort. On the other hand, an expert generally wants to design new algorithms, to integrate new elements into an existing method, or even to combine different search mechanisms. Moreover, such a tool is of large interest in order to be able to evaluate and to compare different algorithms fairly.

Hence, as pointed out in Cahon et al. (2004) and Talbi (2009), three major approaches exist for the development of metaheuristics: from scratch or no reuse, code reuse only and both design and code reuse. Firstly, programmers are tempted to develop and implement their own code from scratch. However, it requires time and energy and the resulting code is generally error prone and difficult to maintain and evolve. The second approach consists of reusing a third-party source code available on the web, either as individual programs or as libraries. Individual programs often have application-dependent sections that are to be extracted before a new application-dependent code is to be inserted. Similarly, modifying these sections is often time consuming and error prone. Code reuse through libraries is obviously better because they are often well tried, tested, documented, and thus more reliable. However, libraries do not allow the reuse of the complete invariant part of the algorithms related to the design. Therefore, the code effort remains important. At last, both design and code reuse allow to overcome this problem. As a consequence, an approved approach for the development of metaheuristics is the use of frameworks.

A metaheuristic software framework may be defined by a set of building blocks based on a strong conceptual separation of the invariant part and the problem-specific part of metaheuristics. Thus, each time a new optimization problem is to be tackled, both code and design can directly be reused in order to redo as little code as possible. Hence, the implementation effort is minimal with regards to the problem under investigation. Generally speaking, the constant part is encapsulated in generic or abstract skeletons that are implemented in the framework. The variable part, which is problem specific, is fixed in the framework but must be supplied by the user. These user-defined functions are thus to be called by the framework. To do so, the design of the framework must be based on a clear conceptual separation between the resolution methods and the problem to be solved. Object-oriented design and programming are generally recommended for such a purpose. But another way to perform this separation is to provide a set of modules for each part, and to make them cooperate through text files. However, this allows less flexibility than the object-oriented approach, and the execution is generally much more time-consuming. Besides, note that two types of software frameworks can be distinguished: white-box and black-box frameworks.

3.2. ParadisEO and ParadisEO-MOEO

ParadisEO (http://paradiseo.gforge.inria.fr) is a white-box object-oriented software framework dedicated to the flexible design of metaheuristics for optimization problems of continuous, discrete and combinatorial nature. Based on EO (evolving objects, http://eodev.sourceforge.net) (Keijzer et al., 2001), this template-based C++ computation library is portable across both Unix-like and Windows systems. This software is governed by the CeCILL license under French law and abiding by the rules of distribution of free software (http://www.cecill.info). ParadisEO tends to be used by both non-specialists and optimization experts. As illustrated in Fig. 1, it is composed of four connected modules that constitute a global framework. Each module is based on a clear conceptual separation of the solution methods from the problems they are in-
tended to solve. This separation confers a maximum code and design reuse to the user. The first module, ParadisEO-EO, provides a broad range of classes for the development of population-based metaheuristics, including evolutionary algorithms and particle swarm optimization techniques. Second, ParadisEO-MO contains a set of tools for single solution based metaheuristics, i.e., local search, simulated annealing, tabu search, etc. Next, ParadisEO-MOEO, which is of our interest in this paper, is specifically dedicated to the reusable design of metaheuristics for multiobjective optimization. Finally, ParadisEO-PEO provides a powerful set of classes for the design of parallel and distributed metaheuristics: parallel evaluation of solutions, parallel evaluation function, island model and cellular model. In the frame of this paper, we will exclusively focus on the module devoted to multiobjective optimization, namely ParadisEO-MOEO.

ParadisEO-MOEO provides a flexible and modular framework for the design of EMO metaheuristics. Its implementation is based on the model presented in the previous section and is conceptually divided into fine-grained elements. On each level of its architecture, a set of abstract classes is proposed and a wide range of instantiable classes, corresponding to different state-of-the-art strategies, are also provided. Moreover, as the framework aims to be extensible, flexible and easily adaptable, all its elements are generic so that its modular architecture allows to quickly and conveniently develop any new scheme with a minimum code writing. The underlying goal here is to follow new strategies coming from the literature and, if need be, to provide any additional algorithmic part required for their implementation. ParadisEO-MOEO constantly evolves and new features might be regularly added to the framework in order to provide efficient and modern concepts and to reflect the most recent advances of the EMO field.

3.3. Main characteristics

A framework is usually intended to be exploited by a large number of users. Its exploitation could only be successful if a range of user criteria are satisfied. Therefore, the main goals of the ParadisEO software framework are the following ones (Cahon et al., 2004; Talbi, 2009):

- **Maximum design and code reuse.** The framework must provide a whole architecture design for the metaheuristic approach to be used. Moreover, the programmer may redo as little code as possible. This aim requires a clear and maximal conceptual separation of the solution methods and the problem to be solved. The user might only write the minimal problem-specific code and the development process might be done in an incremental way, so that it will considerably simplify the implementation and reduce the development time and cost.
- **Flexibility and adaptability.** It must be possible to easily add new features or to modify existing ones without involving other algorithmic elements. Users must have access to source code and use inheritance or specialization concepts of object-oriented programming to derive new objects from base or abstract classes. Furthermore, as existing problems evolve and new others arise, the framework must be conveniently specialized and adapted.
- **Utility.** The framework must cover a broad range of metaheuristics, problems, parallel and distributed models, hybridization mechanisms, etc. Of course, advanced features must not add any difficulty for users wanting to implement classical algorithms.
- **Transparent and easy access to performance and robustness.** As the optimization applications are often time consuming, the performance issue is crucial. Parallelism and distribution are two important ways to achieve high performance execution. Moreover, the execution of the algorithms must be robust in order to guarantee the reliability and the quality of the results. Hybridization mechanisms generally allow to obtain robust and better solutions.
- **Portability.** In order to satisfy a large number of users, the framework must support many material architectures (sequential, parallel, distributed) and their associated operating systems (Windows, Linux, MacOS).
- **Easy-of-use and efficiency.** The framework must be easy to use and must not contain any additional cost in terms of time or space complexity in order to keep the efficiency of a special-purpose implementation. On the contrary, the framework is intended to be less error prone than a specifically developed metaheuristic.

The ParadisEO platform honors all the above-mentioned criteria and aims to be used by both non-specialists and optimization experts. Furthermore, the ParadisEO-MOEO module must cover additional goals related to EMO. Thus, in terms of design, it might for instance be a commonplace to extend a single-objective optimization problem to the multiobjective case without modifying the whole metaheuristic implementation.

3.4. Existing software frameworks for evolutionary multiobjective optimization

Many frameworks dedicated to the design of metaheuristics have been proposed so far. However, very few are able to handle MOPs, even if some of them provide a few particular EMO strategies, such as ECJ [http://cs.gmu.edu/eclab/projects/ecj/], JavaEVA (Streichert and Ulmer, 2005) and Open BEAGLE (Gagné and Parizeau, 2006). Table 2 gives a non-exhaustive comparison between a number of existing software frameworks for EMO, including jMetal (Durillo et al., 2006), the MOEA toolbox for Matlab (Tan et al., 2001), MOMHLib++ [http://home.gna.org/momh/], PISA (Bleuler et al., 2003) and Shark (Igel et al., 2008). Note that other software packages exist for multiobjective optimization (Poles et al., 2008), but some cannot be considered as frameworks and others do not deal with EMO. The frameworks presented in Table 2 are distinguished according to the following criteria: the kind of MOPs they are able to tackle (continuous and/or combinatorial problems), the availability of statistical tools (including performance metrics),

<table>
<thead>
<tr>
<th>Framework</th>
<th>Problems</th>
<th>Statistical tools</th>
<th>Hybrid.</th>
<th>Parallel</th>
<th>Type</th>
<th>Lang.</th>
<th>License</th>
</tr>
</thead>
<tbody>
<tr>
<td>jMetal</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>White</td>
<td>Java</td>
<td>Free</td>
</tr>
<tr>
<td>MOEA for Matlab</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Black</td>
<td>Matlab</td>
<td>Free/com.</td>
</tr>
<tr>
<td>MOMHLib++</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>White</td>
<td>C++</td>
<td>Free</td>
</tr>
<tr>
<td>PSA</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Black</td>
<td>Any</td>
<td>Free</td>
</tr>
<tr>
<td>Shark</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>White</td>
<td>C++</td>
<td>Free</td>
</tr>
<tr>
<td>ParadisEO</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>White</td>
<td>C++</td>
<td>Free</td>
</tr>
</tbody>
</table>

the availability of hybridization or parallel features, the framework type (black-box or white-box), the programming language and the license type (free or commercial).

Firstly, let us mention that every listed software framework is free of use, except the MOEA toolbox designed for the commercial-software Matlab. They can all handle continuous problem, but only a subpart is able to deal with combinatorial MOPs. Moreover, some cannot be considered as white-box frameworks since their architecture is not decomposed into objects. For instance, to design a new algorithm under PSA, it is necessary to implement it from scratch, as no existing element can be reused. Similarly, even if Shark can be considered as a white-box framework, its building blocks are not as fine grained as the ones of ParadisEO. On the contrary, ParadisEO is an open platform where anyone can contribute and add his/her own features. Finally, only a few ones are able to deal with hybrid and parallel metaheuristics at the same time. Hence, with regards to the taxonomy proposed by Talbi (2002), only relay hybrid metaheuristics can be easily implemented within JMetal, MOMHLib++ and Shark, whereas ParadisEO provides tools for the design of all classes of hybrid models, including teamwork hybridization. Furthermore, in opposition to JMetal and MOMHLib++, ParadisEO offers easy-to-use models for the design of parallel and distributed EMO algorithms. Therefore, ParadisEO seems to be the only existing software framework that achieves all the aforementioned goals.

3.5. Implementation

Technical details on the implementation of EMO algorithms under ParadisEO-MOEO can be found on the ParadisEO website (http://paradiseo.gforge.inria.fr), together with a complete documentation and many examples of use. The high flexibility of the framework and its modular architecture based on the three main multiobjective metaheuristic design issues (fitness assignment, diversity preservation and elitism) allow to implement efficient algorithms in solving a large diversity of MOPs. The granular decomposition of ParadisEO-MOEO is based on the conceptual model introduced in the previous section.

3.5.1. EMO algorithms

An EMO algorithm can easily be designed using the fine-grained building blocks of ParadisEO. Different operators can be experimented without engendering significant modifications in terms of code writing. A wide range of strategies are already provided, but this list is not exhaustive as the framework perpetually evolves and offers all that is necessary to develop new ones with a minimum effort. Indeed, ParadisEO is a white-box framework that tends to be flexible while being as user-friendly as possible.

In order to satisfy both the beginners and the more experienced users, ParadisEO-MOEO also provides even more easy-to-use EMO algorithms. These classes propose different implementations of some state-of-the-art methods. They are based on a simple combination of building blocks, as described in Section 2.2. Hence, MOGA (Fonseca and Fleming, 1993), NSGA (Srinivas and Deb, 1994), NSGA-II (Deb et al., 2002), SPEA2 (Zitzler et al., 2001), IBEA (Zitzler and Künzli, 2004) and SEEA (Lieufooghe et al., 2010) are proposed in a way that a minimum number of problem- or algorithm-specific parameters are required. For instance, to instantiate NSGA-II for a new continuous MOP, it is possible to use standard operators for representation, initialization and variation, so that the evaluation function is the single part to be implemented. These easy-to-use algorithms also tend to be used as references for a fair performance comparison in the academic world, even if they are also well suited for a straight use to solve real-world MOPs.

3.5.2. Parallel EMO algorithms

The design and the implementation of sequential EMO algorithms is just an aspect of the features provided by ParadisEO. Indeed, EMO algorithms can easily be hybridized, parallelized and distributed, thanks to the PEO module of ParadisEO (Cahon et al., 2004). The framework has been designed in such a way that the user can, for instance, parallelize his/her sequential EMO algorithm very easily and transparently. When solving real-world optimization problems, running simple EMO algorithms on a large population or on complex individuals often requires high computational resources. However, since each population member can be seen as an independent unit, parallelism naturally arises while dealing with population-based metaheuristics. Melab et al. (2006) identify three levels of parallelism for evolutionary computation, and then EMO algorithms: the solution level, the iteration level and the algorithmic level. ParadisEO is one of the rare framework that provides these most common parallel and distributed models. The resulting applications are portable on different execution platforms such as parallel computing, cluster computing, Internet computing and grid computing.

Evaluating a solution in the objective space is by far the most computationally expensive step of any metaheuristic for difficult or large-size optimization problems. Firstly, at the solution level, the evaluation of a single solution can be parallelized, either partitioning data or objective functions. However, such a parallel evaluation function strongly relies on the problem to be solved so that the user must have enough knowledge about parallel computing to implement it efficiently. Second, at the iteration level, the evaluation of the evolving population can be parallelized as well. At each algorithm iteration, the offspring solutions are distributed to different workers in order to be evaluated in parallel. In terms of implementation, the latter strategy can be done by embedding the original, either sequential or parallel, evaluation function into a more advanced object with which the EMO algorithm is to be created. Note that these two first parallelization schemes do not modify the behavior of the EMO algorithm but are essentially focused on speeding up the search. At last, the algorithmic level consists of running simultaneously a number of EMO algorithms in parallel. This can be done in a multi-start way, what seems to be particularly interesting for testing different independent algorithms or parameter settings. However, various EMO algorithms can also cooperate with each other during the search process, either in a centralized way or like in the island model. The approach results in a hybrid co-evolutionary EMO algorithm where different agents cooperate in a high-level teamwork mode (Talbi, 2002). For instance, in the island model, sub-populations are distributed in a set of islands in which semi-isolated EMO algorithms are executed in parallel. Each island manages a proper population and an optional proper archive. However, a global archive containing the set of non-dominated solutions from all the islands can possibly be designed as well. The migration of individuals from an island to another is managed by a migration decision condition, an exchange topology, a number of emigrants, an emigrants selection scheme and an integration strategy. Note that emigrant individuals can be selected from both the current population and the current archive of a given island. The reader is referred to (Cahon et al., 2004) for more information about the ParadisEO parallel models.

3.6. Discussion

We believe that the aforementioned characteristics make ParadisEO-MOEO a valuable tool for both researchers and practitioners, and a unique software framework in comparison to existing ones. Indeed, it includes many state-of-the-art and up-to-date EMO algorithms. The rich set of ParadisEO modular ingredients

serve as building blocks to implement these methods. The related source code is maintained and regularly updated by the developers. Moreover, the framework gives the possibility to design and implement a wide number of new MOP resolution methods, either sequential or parallel, just by combining existing elements in an innovative way, or by implementing original ones. As an example, to illustrate the high ParadisEO flexibility, the dominance relation used to manage the archive of a given EMO algorithm is just a parameter of the corresponding object. The latter can then be modified from, say, Pareto-dominance to ϵ-dominance with insignificant effort in order to obtain a different method. Furthermore, analogously to the design level, a large number of issues are shared by many EMO algorithms at the implementation level as well. Hence, ParadisEO can serve as a reference implementation in order to compare different algorithms fairly. For instance, whenever a new EMO algorithm is proposed, its efficiency can be experimentally demonstrated by comparing its behavior with existing ones.

On the other hand, ParadisEO is also a practical tool that can be used to tackle an original MOP. The implementation of efficient programs is highly facilitated so that the user only has to focus on problem-related issues of representation, initialization, evaluation and variation. The implementation effort is even more reduced when a classical solution representation can be applied for the problem under consideration. By means of classical representation, we include real-coded, binary, integer, permutation variable representations, and so on. For such problems, the development and time cost are reduced to minimum since the evaluation function is the single element to be implemented. Of course, this cost is always related to the proficiency of the programmer in charge of the implementation. Once this evaluation function is available, the user only has to instantiate any EMO algorithm (NSGA-II, SPEA2, IBEA, …) and optionally any parallel model (parallel evaluation, island model, …) to obtain a powerful resolution program that is able to run on a large range of material architectures (sequential, parallel, distributed) and their associated operating systems (Windows, Linux, MacOS). Therefore, the remaining development consists of a direct instantiation of selected strategies. Though, for more sophisticated solution encodings, the development cost remains substantial with respect to the complexity of the underlying representation and to the level of expertise of the programmer. But it will always be lower than implementing a whole specific EMO algorithm from scratch. Otherwise, starting from a single-objective optimization problem implemented within ParadisEO, it is a commonplace to investigate a multiobjective variant, so that existing EMO algorithms can then be instantiated to solve the resulting MOP.

With regards to runtime efficiency, it is difficult, if not impracticable, to show how ParadisEO-MOEO performs in comparison to existing software frameworks or even to a specific implementation done from scratch. Indeed, it will always depend on the programmer ability to handle such a tool, and may often depend on the problem to be solved and the algorithm to be applied. Development from scratch generally allows a deep code optimization with respect to a very specific application, but this requires a high level of expertise and is also time-consuming in terms of development cost. Additionally, ParadisEO-MOEO has a clear advantage of not communicating information through text files, as done in PISA for instance, what is known to be very expensive in time.

Finally, ParadisEO-MOEO has been used and experimented to solve a large range of MOPs from both academic and real-world fields, which validates its high flexibility. Indeed, various academic MOPs have been tackled, including continuous test functions such as the ZDT and DTLZ functions family (Deb et al., 2005b) or the Schaffer problem (Schaffer, 1985), scheduling problems including permutation flow-shop scheduling (Liesfooghe et al., 2007a), routing problems such as the multiobjective traveling salesman problem and the bi-objective ring star problem (Liefooghe et al., 2010). Moreover, ParadisEO-MOEO has been successfully applied to solve real-world applications in structural biology (Boisson et al., 2008), feature selection in cancer classification (Talbi et al., 2008), materials design in chemistry (Schuetze et al., 2008), portfolio optimization, etc. Note that the problem-related part of some aforementioned applications is freely available as contributions on the ParadisEO website (http://paradiseo.gforge.inria.fr), together with a detailed documentation and some tutorial lessons.

Additionally, hybrid and parallel metaheuristics have also been designed within ParadisEO to solve MOPs. For instance, hybrid EMO algorithms have been experimented by Liefooghe et al. (2010), a multiobjective cooperative island model has been designed by Talbi et al. (2007), costly evaluation functions have been parallelized by Boisson et al. (2008), and a parallel multiple reference point approach has been proposed by Figueira et al. (2010).

4. Concluding remarks

The current paper described a software framework for the development of evolutionary multiobjective optimization algorithms. First, we identified the common concepts shared by many evolutionary multiobjective optimization techniques, separating the problem-specific part from the invariant part of such approaches. We emphasized the main issues of fitness assignment, diversity preservation and elitism. Therefore, we presented a conceptual model, based on a fine-grained decomposition, and we illustrated its robustness and reliability by treating a number of state-of-the-art algorithms as simple instances of it. Next, this unified view has been used as a starting point for the design and the implementation of a general-purpose software framework called ParadisEO-MOEO. ParadisEO-MOEO is a free C++ white-box object-oriented framework dedicated to the flexible and reusable design of evolutionary multiobjective optimization algorithms. It is based on a clear conceptual separation between the resolution methods and the problem they are intended to solve, thus conferring a maximum code and design reuse. This global framework has been experimentally validated by solving a comprehensive number of both academic and real-world multiobjective optimization problems.

However, we believe that a large number of issues from evolutionary multiobjective optimization are shared by many other metaheuristic methodologies. Thereafter, we plan to generalize the conceptual model introduced in this paper to other existing metaheuristic approaches for multiobjective optimization. Hence, multiobjective local search or scatter search methods might be interesting paths to explore in order to investigate their ability and their modularity for providing such a flexible model as the one presented in this paper. The resulting general-purpose models and their particular mechanisms would then be integrated into the ParadisEO-MOEO software framework.

Acknowledgments

This work was supported by the ANR DOCK project. The authors would like to gratefully acknowledge Jérémie Humeau, Thomas Legrand, and Abdel-Hakim Deneche for their helpful contribution on the implementation part of this work, as well as Sébastien Cahon and Nouredine Melab for their work on the preliminary version of the ParadisEO-MOEO software framework presented in this paper. Moreover, the authors would like to thank the anonymous reviewers for their valuable comments and suggestions.
References


