Multiobjective Optimization in Bioinformatics and Computational Biology

Julia Handl, Douglas B. Kell, and Joshua Knowles

Abstract—This paper reviews the application of multiobjective optimization in the fields of bioinformatics and computational biology. A survey of existing work, organized by application area, forms the main body of the review, following an introduction to the key concepts in multiobjective optimization. An original contribution of the review is the identification of five distinct "contexts," giving rise to multiple objectives: These are used to explain the reasons behind the use of multiobjective optimization in each application area and also to point the way to potential future uses of the technique.

Index Terms—Global optimization, clustering, classification and association rules, interactive data exploration and discovery, experimental design, machine learning, bioinformatics (genome or protein) databases.

1 Introduction

Numerous problems encountered in bioinformatics and computational biology can be formulated as optimization problems and, thus, lend themselves to the application of powerful heuristic search techniques [1], [2]. Traditionally, the optimization is conducted with respect to a single "goal," but the possibility of optimizing *multiple* objectives *simultaneously* is rapidly becoming more recognized [3], [4], [5], [6], [7], [8], [9], [10], [11]. Recently, in biology, multiobjective optimization has been shown to have significant benefits compared to single-objective approaches, e.g., in classification [12], system optimization [13], [14], and inverse problems [15].

In this paper, we aim to outline the potential scope of methods for multiobjective optimization in biological applications and to provide a review of existing work. After a brief reminder of the basic concepts of multiobjective optimization (Section 2), we proceed in Section 3 to identify five distinct contexts in which multiple objectives may arise, or be used, in solving an optimization problem. Sections 4, 5, 6, 7, and 8 contain the bulk of the survey material, organized by application area. References to the earlier categorization by context are made in these sections in order to achieve our principal aim in this review: to unravel the variety of motivations behind the uses of multiobjective optimization in biological applications. Section 9 discusses our findings and issues arising from the survey, while Section 10 concludes.

2 Multiobjective Optimization

Multiobjective optimization (MOO) concerns optimization problems with multiple objectives (a.k.a. goals or criteria). Typically, the objectives may estimate very different aspects of the solutions, aspects that are, therefore, *incommensurable* and often (partially or wholly) in conflict.

A general (unconstrained) multiobjective optimization problem (MOP) can be defined mathematically as:

$$\begin{aligned} &''\text{minimize}'' \ \mathbf{z} = \mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_m(\mathbf{x})) \\ &\text{with } \mathbf{x} = (x_1, x_2, \ldots, x_n) \in X, \end{aligned} \tag{1}$$

where $\mathbf{x}$ is an n-dimensional decision vector or solution and X is the decision space, i.e., the set of all expressible solutions. The vector objective function $\mathbf{f}(\mathbf{x})$ maps X into $\mathbb{R}^m$, where $m \geq 2$ is the number of objectives. The vector $\mathbf{z} = \mathbf{f}(\mathbf{x})$ is an objective vector or point. The image of X in objective space is the set of all attainable points, Z (see Fig. 1).

The term "minimize" appears above in quotation marks because its meaning is not yet defined. Alternative minimization problems exist, including lexicographic optimization (e.g., as used in Olympic games medal tables), minimizing the maximum of all the objectives (minmax), and minimizing a scalarized combination of the objectives (see [16], [17]). However, by far the most frequent understanding of "minimize," above, is in the sense of Pareto optimality. The *Pareto optimal set* X^* of solutions consists of all those that it is impossible to improve in any objective without a simultaneous worsening in some other objective:

$$\begin{aligned} &X^* = \{\mathbf{x}^* \in X \mid \nexists \mathbf{x} \in X, \quad \mathbf{f}(\mathbf{x}) \leq \mathbf{f}(\mathbf{x}^*)\}, \text{where } \mathbf{f}(\mathbf{x}^1) \\ &\leq \mathbf{f}(\mathbf{x}^2) \text{ iff } \forall i \in 1..m, f_i(\mathbf{x}^1) \leq f_i(\mathbf{x}^2) \wedge \exists j \in 1..m, \\ &f_j(\mathbf{x}^1) < f_j(\mathbf{x}^2). \end{aligned} \tag{2}$$

The points in objective space corresponding to the Pareto optima are termed *nondominated* and form the *Pareto front*.

In most cases, the Pareto optimal set contains more than one element because there exist different trade-off solutions to the problem which offer different compromises of the objectives. Thus, in practice, solving an MOP often means that a human decision-maker (DM) is involved who then chooses a solution that is Pareto optimal (ideally). Methods

• *The authors are with the School of Chemistry, The University of Manchester, Faraday Building, PO Box 88, Sackville Street, Manchester M60 1QD, UK.*
E-mail: {j.handl, dbk, j.knowles}@manchester.ac.uk.

Manuscript received 1 Mar. 2006; accepted 1 June 2006; published online 12 Jan. 2007.
For information on obtaining reprints of this article, please send e-mail to: tcbb@computer.org, and reference IEEECS Log Number TCBBSI-0034-0306. Digital Object Identifier no. 10.1109/TCBB.2007.070203.

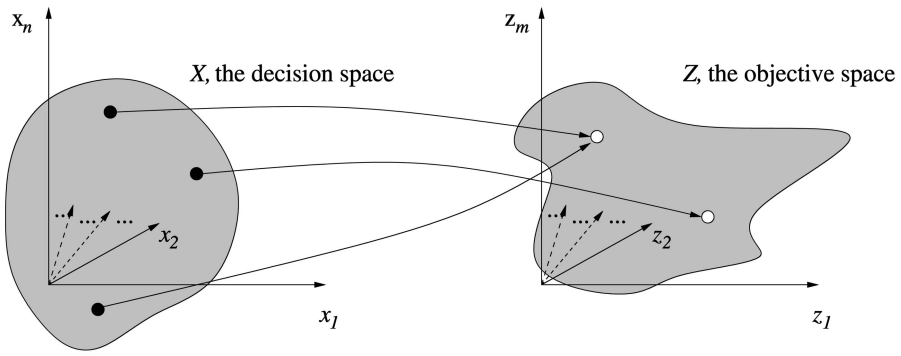

Fig. 1. The n-dimensional parameter space maps to the m-dimensional objective space.

of decision-making (before, after, or interactively during search) have been extensively investigated in a branch of management science/operations research known as multicriterion decision making (MCDM) [18], [19] and may include advanced methods of visualization, e.g., [20].

The a posteriori and interactive methods of decision-making can be more effective as the decision-maker may be helped by "seeing" what trade-off solutions are possible. This view has led to a burgeoning of methods for generating the whole Pareto set or an approximation to it.

Several distinct types of methods for generating good approximations to the Pareto set have been developed, e.g., see [3], [4]. Evolutionary algorithm approaches have become particularly popular and good overviews of these can be found in [5], [6], [7], [8]. For the use of multiobjective optimization techniques in specific application domains, see, e.g., [9], [10], [11].

Note that, in the remainder of this review, when we use the term "multiobjective optimization" or "MOO," we refer by default to the process of generating the whole Pareto set/front or an approximation to it; when we refer to other methods of tackling MOPs (e.g., combining objectives into a scalar function), we point these out explicitly.

3 Five Distinct Contexts Giving Rise to Multiple Objectives

In this review paper, our principal aim is to unravel the different *reasons* underlying the need for multiobjective optimization in biological applications. To achieve this, we propose here a categorization based on the different types of contexts in which multiple objectives may arise or be usefully exploited. The literature reviewed in the survey remains arranged by biological problem domain, but a classification with respect to the five contexts introduced here is provided at the end of this review, in Section 9.

3.1 Standard MOO

As a first category, we identify the "standard" context of multiobjective optimization, where all objectives are clear, measurable goals that we would genuinely like to optimize. Assuming all important criteria have been included as objectives, we may be unsure about their relative importance but we are certain that our "ideal" solution will be Pareto optimal. Thus, using an approach that generates a Pareto front (approximation), a decision maker can learn something about the conflicts between the objectives, the space of possible solutions, and may subsequently select a single preferred solution.

An example of this type of problem setting is the optimization of biochemical processes where trade-offs exist between aspects of product quality and reaction time or throughput (see Section 8.2).

3.2 Counterbalance for Bias

The second category is where MOO is used as a tool to counterbalance a measurement bias affecting an objective function. Such a measurement bias is, for example, encountered in alignment problems, where short alignments can be trivially obtained and the number of mismatches automatically increases with the length of the alignment.

Mathematically, this setting can be described as follows, assuming just one (primary) objective to be optimized:

$$f(\mathbf{x}) = f'(\mathbf{x}) + m(g(\mathbf{x})), \tag{3}$$

where f' is an ideal (i.e., unknown), unbiased measure of the primary objective, $m(g(\mathbf{x}))$ is a bias term where m is an unknown but monotone function of a measurable function g, and f is the measurable but biased sum of the two. In the example of alignment problems, f (the scoring function used) gives a (biased) quality estimate, g is the length of the given alignment, m is assumed to be a monotone function, and f' is the ideal (but unknown) quality of the alignment.

We would like to minimize $f'(\mathbf{x})$ as follows:

$$\text{minimize } f'(\mathbf{x}) = f(\mathbf{x}) - m(g(\mathbf{x})), \tag{4}$$

but, since m is unknown, we cannot formulate the problem in this way. However, we may formulate the problem instead as:

$$\begin{aligned} &''\text{minimize}'' \ (f(\mathbf{x}), -(g(\mathbf{x}))), \\ &\text{with } \mathbf{x} = (x_1, x_2, \ldots, x_n) \in X, \end{aligned} \tag{5}$$

in terms of two measurable objectives. Hence, the framework of MOO is used as a means of introducing an additional objective, g, to counterbalance the bias of the primary objective.[1]

The set of Pareto optimal solutions will certainly contain the desired solution since each Pareto optimum is the best

1. N.B. the equations above can be generalized to more than one primary objective, where necessary.

value of $f(x)$, given a fixed value of $g(x)$. In this scenario, selection of the best solution does not usually depend on preferences, but on the estimation of the biases. In some applications, the biases may be estimated using random control data and this may help to identify the best solution in the Pareto front.

Examples of this type of problem include unsupervised feature selection and sequence and structure alignment problems (see Sections 4.2.2 and 6).

3.3 Multiple Source Integration

In the third category, MOO is used to integrate noisy data from multiple sources. Hence, in this setting, it is used as an alternative to an a priori or a posteriori integration technique. The problems where this approach is used are often originally single-objective. However, multiple noisy views of the data need to be integrated as their combined use may yield better results than the use of data from a single information source.

Mathematically, this setting can be described by a set of objective functions:

$$\begin{aligned} f_1(\mathbf{x}) &= f'_1(\mathbf{x}) + \bar{n}_1 \\ &\vdots \\ f_m(\mathbf{x}) &= f'_m(\mathbf{x}) + \bar{n}_m, \end{aligned} \tag{6}$$

where the function value of each objective function f_i is equal to the value of an ideal function f'_i with some unknown random noise $\bar{n}_i$ on it, for $i \in 1..m$. In some cases, the f' are all identical, i.e., the "views" of the data arise from the *same* types of measurement but, e.g., taken at different times. By formulating the problem as

$$\begin{aligned} &''\text{minimize}'' \ \mathbf{z} = \mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_m(\mathbf{x})) \\ &\text{with } \mathbf{x} = (x_1, x_2, \ldots, x_n) \in X \end{aligned} \tag{7}$$

and finding the Pareto optima, the impact of the noise may be reduced if it is reasonably uncorrelated with the solution space X. Nonetheless, note that it is not guaranteed that the desired solution will be among the Pareto optima.

Examples of this type of problem are the inference of phylogenetic trees and data clustering with several dissimilarity matrices (see Sections 5.1 and 4.2.1).

3.4 Performance Approximation by Proxies

Category four is comprised of those applications in which the "real," underlying objective of the problem, $f'(\mathbf{x}, \mathbf{y})$, is a function of both the solution $\mathbf{x}$ and some "hidden" variables $\mathbf{y}$ that are not available during optimization. For example, in training a supervised classifier, $\mathbf{y}$ refers to the generalization ability of the classifier on *future* data (which may be estimated using a test set *after* the optimization, but the classifier must not be trained using these examples).

Since the function f' is not suitable for use in the optimization process (because $\mathbf{y}$ is unavailable), it needs to be replaced by "proxy" objectives $f_i(\mathbf{x})$, which are functions of $\mathbf{x}$ only. Often, such "proxy" objectives only capture certain aspects of a good solution and different proxies are complementary with respect to each other. Thus, it should be expected that the desired solution(s) will score relatively highly under all of the "proxy" objectives and a MOO approach therefore seems useful, although the desired solution cannot be *guaranteed* to be among the associated set of Pareto optima.

Note that the difference between this context and that of standard MOO (as introduced above) may seem unclear to some readers. However, the distinction is clear: In the case of standard MOO, the objective functions have primacy, i.e., it is they that define the Pareto set, e.g.: The concept of a "best car" does not exist *per se*, but, given a search space, a set of "best cars" is *induced* by the objectives chosen. In contrast, in the context of proxy objectives, it is the solution that has primacy and the objectives are only a means of orienting the search in order to discover this solution, e.g., the real structure of a protein exists and we may try and find it by employing a number of different energy/cost functions.

Examples of this type of problem include supervised classifier training (as explained above), data clustering, and protein structure prediction (see Sections 4.1, 4.2.1, and 7).

3.5 Multiobjectivization

The fifth and final category we identify is where MOO may be used solely as a way to obtain improved search "guidance" in what is essentially a single-objective problem.[2] Assuming a single objective that is measurable, a problem may still be difficult because of its search landscape. There are at least two difficulties in search landscapes that can potentially be reduced by "multiobjectivization": 1) where a problem involves frustration (or epistasis), which causes excessive local optima in the search landscape and 2) where the search landscape contains regions offering little or no objective function gradient. In the first case, decomposition of the primary objective into several different functions (each function either defined over all of the variables or a subset of them) may help to separate out the conflicting aspects of the problem, thus reducing the number of local optima "seen" by a search algorithm [21]. In the second case, the use of extra "helper objectives" in addition to the primary objective may provide helpful guidance in the flat regions of the landscape [22], [21].

Multiobjectivization may potentially be achieved by any reformulation of the problem that respects the following relation [21]:

$$\forall \mathbf{x}^{opt} \in X \ \exists \mathbf{x}^* \in X, x^* = x^{opt}, \tag{8}$$

where $\mathbf{x}^{opt}$ is an optimal solution to the original single-objective problem and $\mathbf{x}^*$ is a Pareto optimum of the multiobjectivized problem. This ensures that at least one of the true Pareto optimal solutions will be optimal with respect to the original primary objective and will correspond to the best solution.

An example of this type of problem is structure identification from X-ray powder diffraction data (see Section 7.1).

2. N.B. there is no reason why multiobjectivization cannot also be generalized to the case where the original problem is multiobjective.

4 Classification

A large number of problems typically encountered in bioinformatics are classification problems. We will now briefly introduce the main types of classification—unsupervised, supervised, and semisupervised—and consider the applicability of multiobjective approaches in these different areas.

4.1 Supervised Classification

Supervised classification techniques require the presence of training data, that is, a (sufficiently large) set of data samples for which the correct classification is known. The aim in supervised classification is to obtain a classifier with good *generalization properties*, i.e., a classifier that performs well on previously unseen test data. Evidently, this goal of supervised classification cannot be measured objectively during the training of a classifier, so "proxy" objectives must be used to estimate expected generalization performance, e.g., through the use of validation test scores.

4.1.1 Receiver Operating Characteristics (ROC) Curve

When considering the performance of binary classifiers (e.g., for the distinction between tumor and healthy tissue), sensitivity and specificity are often seen as more informative measures of the classification performance than the overall classification accuracy. The sensitivity and specificity of a classifier are always conflicting and, for a given classifier, the trade-off between the two can be represented in the form of a receiver operating characteristics (ROC) curve [23]. Traditionally, this trade-off curve has not been explicitly optimized and, instead, an a priori weighting of the two objectives has been used during training. For the classifier obtained for a given data set, an ROC curve can then be generated by varying one or more of the parameters of the classifier and plotting the impact on the values of sensitivity and specificity. A practitioner could then pick any point on the ROC curve that corresponds to the desired sensitivity or specificity. However, an ROC curve obtained in this way is unlikely to be optimal in the Pareto optimal sense and more favorable trade-offs between sensitivity and specificity can be obtained through the direct use of MOO [24]. ROC curves for multiclass problems have also been optimized using multiobjective evolutionary optimization [25].

4.1.2 Partial Classification/Rule Mining

In partial classification (also referred to as nugget mining or classification rule mining), the aim is to identify and describe interesting subsets of the data only. This has many applications in data mining, where it may be (small) subsets of the data (which show exceptional and/or unexpected behavior) that are of real interest, e.g., in the analysis of gene expression data [26]. Among the most established performance measures in partial classification are the measures of coverage and confidence [27], which reflect the principles of sensitivity and specificity on a local basis. Coverage gives the proportion of total members of a class correctly described by a given rule (i.e., maximizing coverage means minimizing the number of false negatives), whereas confidence gives the proportion of the patterns to which a given rule has been correctly applied (i.e., maximizing confidence means minimizing the number of false positives). Evidently, coverage and confidence are conflicting criteria and multiobjective evolutionary algorithms have therefore been applied to their optimization in rule mining [28], [29]. This is a flexible alternative to the use of fixed thresholds on coverage and/or confidence during or after the optimization process [30].

4.1.3 Balancing Model Accuracy and Complexity

A persistent problem in supervised classification is the trade-off between model performance and model complexity: If sufficiently trained, many types of classifiers can obtain a very high classification accuracy on the training data but may not generalize well subsequently. In general, simple models are therefore preferred during classification (this is also know as Occam's Razor [31]) in order to avoid overtraining and this principle needs to be integrated into the optimization process. Traditional approaches to avoid overtraining include cross-validation, early stopping in the training of neural networks [32], or the pruning of decision trees [33]. In contrast to these, multiobjective optimization provides a more general and flexible framework to integrate model complexity into the optimization process [34]. The benefits of integrating model complexity as a second and/or third objective have been previously demonstrated with respect to fuzzy rule mining [35], learning classifier systems [36], decision trees [37], support vector machines [38], genetic programming (GP) [39], [40], [41], and artificial neural networks [42], [43], [44], [45]. A specific application in bioinformatics has been the use of multiobjective genetic programming for the identification of quantitative structureactivity relationships (QSAR) [41], where model complexity was measured using a number of different aspects, including the total number of terms, the number of nonlinear terms, and a knowledge-based measure of the chemical interpretability of the descriptors used.

4.1.4 Supervised Feature Selection

The problem of supervised feature selection is another example in which the trade-off between classification accuracy and model complexity is relevant. Given two feature sets of different cardinality that result in the same classification accuracy, the smaller of the two is expected to result in a better generalization performance [46]. Multiobjective formulations of the supervised feature selection problem have been proposed which directly capture this intuition [12], [47], [48], [49], [50], [51] and interesting results have been obtained. For example, the analysis in [12] revealed that certain subsets of features are present in many of the feature subsets in the Pareto front and that an analysis of the ensemble of feature sets present in the Pareto front may, therefore, yield novel insight regarding the relative importance of individual features.

A simpler and very common approach to supervised feature selection is the selection of variables based on their discriminatory power with respect to the target classes. This can, for example, be established using statistical t-tests [52], but, for high-dimensional data, as are typically encountered in postgenomic data analysis [53], the approach inherently suffers from multiple testing issues (i.e., a large number of false positives will result due to random correlations).

Alternative objectives regarding the properties of interesting genes may exist, such as small intracluster dispersion, large intercluster dispersion, or absolute differences between expression levels. The use of multiple, uncorrelated criteria for gene selection may help to reduce multiple testing issues. In recent work by Hero et al. [54], [55], the use of a multiobjective formulation was therefore suggested as a formalized way for the screening of genes in the presence of more than one criterion.

4.1.5 Ensemble Learning

The potential of MOO to furnish a decision maker with a choice of trade-off single classifiers from which she may choose one has been dealt with before; an equally attractive proposition is to use the trade-off set to make an ensemble classifier (i.e., a classifier, which bases its output on a voting between the classifiers in the ensemble).

The integration of several diverse classifiers by means of ensemble techniques may prevent overfitting and increase both classification accuracy and robustness of the individual approaches, as well as the level of confidence in the results returned [46], [56]. Diversity between the classifiers may be obtained through the use of conceptually different classifiers, training on different bootstrapping data sets, a reweighting of the input data, or the introduction of noise.

In [57], the use of multiobjective optimization was explored as a tool to simultaneously optimize the classification accuracy on two different training sets and, thus, ensure diversity in the resulting set of Pareto optimal classifiers. An ensemble based on these classifiers was found to produce results comparable to a well-known technique for ensemble creation, negative correlation learning [49]. An alternative multiobjective formulation which explicitly optimizes classification accuracy and diversity of the members of the ensemble was suggested in [58].

4.2 Unsupervised Classification

Unsupervised classification works in the absence of any training data as such, i.e., without knowledge of the class memberships of individual samples. It therefore relies on the presence of distinct structure in the data and it must be hoped that a distance measure or a reduced feature space can be identified under which related data items cluster together in data space. The overall aim in unsupervised classification is to identify interesting patterns in the data. This concept of interestingness is even harder to quantify than that of "generalization performance" in supervised classification and, equally, calls for the use of "proxy" objectives. The potential of multiobjective approaches in this area has been little explored to date; in the following, we summarize these few examples and consider promising areas for future study.

4.2.1 Clustering

Clustering is the partitioning of data into subgroups and is one of the fundamental tasks in unsupervised classification. Many different formulations of the clustering problem exist, the best known of which are based on minimizing intracluster variance [59]. It is well-known that none of the existing clustering criteria can capture all of the different aspects that humans perceive as properties of a good clustering, such as the compactness of clusters, spatial separation between them, and compliance with local density distributions [53]. One possibility to reduce the problem of the failure of a clustering algorithm in scenarios where the clustering criterion employed is inappropriate is the use of ensemble techniques to integrate the results of a variety of different clustering methods [60], [61]. An alternative to this a posteriori integration of different clustering results is the direct optimization of a partitioning with respect to a number of complementary clustering criteria. Recent work has shown that such multiobjective approaches to clustering can indeed result in an improved and robust performance across data exhibiting a range of different data properties and may be superior to some a posteriori integration approaches [62]. This work also illustrated that good clustering solutions tend to give rise to distinct "knees" in the Pareto front and may be automatically identified through a comparison to random control data [62].

The use of multiobjective optimization for clustering has also been proposed for situations in which the clustering criterion is biased with respect to the number of clusters [63] or where multiple sources of data—in the form of multiple dissimilarity matrices—should be integrated into a single clustering [64], [65]. Such data may be tackled through 1) an a priori fusion of the data and the use of a standard clustering algorithm, 2) the use of ensemble techniques for the a posteriori fusion of the different partitionings obtained, or 3) the selection of a primary clustering objective and the definition of all others as constraints in a constrained optimization problem. However, some work [64], [65] argues that a MOO approach may provide more information and choice to a decision maker.

4.2.2 Unsupervised Feature Selection

Feature selection as part of clustering can be beneficial as the input data may contain many noisy or irrelevant variables, which will hide the structure in the data. It is therefore important to develop algorithms that can reduce the set of input variables to those that contain clear cluster structures and may, therefore, be interesting to analyze; this can be achieved either through specialized clustering algorithms, which explicitly search feature subspaces (such as biclustering algorithms, see [66]), or algorithms for feature selection that can be used as preprocessing methods for the subsequent application of any clustering method.

While the subject of supervised feature selection has been thoroughly explored in the literature, little work exists on the topic of unsupervised feature selection. This is due to the difficulty of the formalization of criteria for the objective assessment of the quality of different feature subspaces. One particular problem is the comparison of feature subspaces of different cardinality as existing measures are usually biased toward small or large feature subspaces [67]. Multiobjective optimization has recently been introduced as a potential solution to this problem as it allows one to optimize one of these objectives and to counterbalance its bias through the simultaneous minimization or maximization of feature cardinality [68], [69], [70].

4.2.3 Association Rule Mining

Association rule mining is the unsupervised equivalent to classification rule mining; thus, it is not pattern-class relationships that are sought, but relationships between patterns in the data space. The quality of association rules is inherently difficult to assess and a range of different objectives have been introduced in the literature [71]. These include measures related to confidence and coverage in classification rule mining, but are also comprised of methods assessing the complexity of the rules found. As several of these measures are complementary and conflicting, their multiobjective optimization has been proposed by a number of authors [72], [73]. For example, in [73], a five-objective formulation of the problem was suggested and a multiobjective evolutionary algorithm was employed for the identification of an optimal set of association rules on a gene expression data set.

4.2.4 Multidimensional Scaling

Next to clustering, rule discovery, and feature selection, another important problem in unsupervised classification is the projection of a data set to lower-dimensional subspaces. Usually, a projection to a two or three-dimensional subspace will be used, with the aim of obtaining a visualization of the data set that is interpretable by a human observer.

Multidimensional scaling (MDS, [74]) is an example of such a visualization technique. Given information about the dissimilarities between data items, MDS provides an embedding of these data into a multidimensional space of specified dimensionality such that distances between data items are preserved, but the actual positions of individual data items are meaningless. A range of different methods for multidimensional scaling exist which differ in the loss function (optimization criterion) used. Brusco [75] suggests that a MOP formulation of the multidimensional scaling problem may be advantageous for two different reasons. A first advantage would be the opportunity to consider a number of different loss functions simultaneously. Alternatively, the multiobjective framework could be used for multidimensional scaling in the presence of multiple dissimilarity matrices, where the same loss function is optimized with respect to the individual dissimilarity matrices.

4.3 Semisupervised Classification

Semisupervised classification combines the techniques of supervised and unsupervised classification (reviewed above) in order to exploit both a (small) number of samples with known classes and a (larger) number of unlabeled data. Data sets with both labeled and unlabeled elements are frequently encountered in application domains where the categorization of individual data items is accompanied by high computational, analytical, or experimental costs, as is the case, e.g., in protein classification [76].

Given the need for the integration of unlabeled and labeled data, as well as a lack of knowledge regarding the importance (weighting) to attach to each and their respective correctness, their simultaneous consideration within a multiobjective framework seems natural. In preliminary work on this subject [77], [78], promising results have been obtained. In particular, multiobjective formulations of both semisupervised clustering and semi-supervised feature selection were shown to outperform single-objective approaches based on the linear and non-linear combination of the objective components related to unlabeled and labeled data.

5 Inverse Problems

Inverse modeling problems arise in all of those applications where the data generated by a biological process or system can be measured and where we are aiming to infer the original system from the observed data [79]. The challenges typically encountered in these applications include noisy data, the integration of several types of data, and the underdetermination of the inference problem at hand. In the following, we will review how multiobjective optimization can serve as a tool to tackle some of these issues.

5.1 Phylogenetic Inference

Phylogenetic tree inference is a special example of an inverse modeling problem encountered in the biochemical domain. Phylogenetic trees, an important tool in evolutionary biology, describe and visualize the evolutionary relationships between species that are believed to have a common ancestor [80]. Existing approaches for phylogenetic tree inference can be classed into three major groups, namely, distance matrix methods, maximum parsimony methods, and maximum likelihood methods [81], [82], [83]. These traditional approaches to phylogenetic tree inference do not take the existence of multiple data sets from different sources into account. While such data sets may often be noisy and partially conflicting, they can usually be assumed to complement each other and to be more informative in combination than on their own. Currently, integration of these different information sources is most commonly done prior to, or after, the actual phylogenetic tree inference [84]. In very recent work, it has been suggested that MOO may provide an alternative tool to integrate and trade-off such conflicting data *during* the inference process and that such an approach may in fact be more robust and informative than the a priori or a posteriori integration currently used in the literature [85].

5.2 Gene Regulatory Networks

Researchers working in the biological sciences today have access to the complete genomes of an increasing number of organisms, but an improved understanding of these organisms requires more knowledge about the activation patterns of the individual genes. Specifically, we need to know and to predict at what level, at what time, under which conditions, and at which location specific genes are expressed in a given organism. The mechanism at the bottom of these patterns of gene expression is a complex interplay of interactions between DNA, RNA, proteins, and metabolites which, at an abstract level, can be modeled as a network of inhibitory and stimulatory interactions between genes—a *gene regulatory network* (GRN).

The actual inference task is typically highly underdetermined, i.e., there may exist (infinitely) many different GRNs that are consistent with the observed data and the

integration of additional information may be necessary to differentiate between these. As discussed in several previous sections, multiobjective optimization can serve as a tool for the flexible integration of data from several sources. In GRN inference, we may additionally use MOO as a means to integrate topological constraints and prior knowledge, which we view as forms of proxy objectives. Some preliminary work [86], [15], [87] has exemplified these uses of MOO for GRN inference and the results obtained seem highly promising.

Clearly, multiobjective approaches are equally relevant when considering related inverse modeling tasks such as time series prediction [88] or the inference of protein networks and metabolic pathways from experimental data [89].

6 Sequence and Structure Alignment

In this section, we will highlight applications related to the assessment of sequential and structural similarities of DNA and RNA macromolecules, as well as proteins. Tools for the assessment of similarity and the identification of related sequences or structures are among the most important tools in bioinformatics as they may serve for the functional categorization of novel genes or proteins of unknown function. This is because sequential and structural similarities may provide evidence of evolutionary relationships between entities and may indicate shared or related functional properties and/or whether evolution is convergent or divergent.

6.1 Sequence Alignment

Sequence alignment aims to arrange two or more DNA, RNA, or protein sequences in a way that highlights their similarities. For this purpose, the sequences are set down one upon the other and are padded with gaps so that a scoring model which assesses the quality of an alignment as a function of the number/type of mismatches, matches, and gaps is optimized [90]. Specifically, the scoring model requires the definition of an appropriate substitution matrix and gap penalty, which indicate the rewards for the alignment of any two characters in the alphabet and the penalty for the insertion of gaps, respectively.

In light of the difficulty of deriving suitable substitution matrices and gap penalties, a multiobjective approach to sequence alignment was proposed in [91]. It is based on a modified form of dynamic programming and, instead of a scalar scoring function (combining all reward and penalty values), a vector of all substitution values and gap penalties is optimized. The same authors have also investigated a biobjective approach to sequence alignment in which only the gap penalties are treated as a separate objective [92]. While a multiobjective formulation may be too expensive to use in certain practical applications, it may certainly prove a useful tool for a thorough analysis of the trade-offs inherent to the problem and for use in the design of novel and improved substitution matrices.

An important distinction in sequence alignment is that between global and local alignments. Global alignment aims to align the sequences given along their entire range. In contrast, local alignment only tries to identify subregions of the sequences in which their patterns concur. As far as local sequence alignment is concerned, a further trade-off can be observed between the length of the patterns compared and the quality scores obtained: Evidently, the number of adverse substitutions or of gap penalties tends to increase for longer alignments, i.e., there is a bias causing the preference of short alignments. In recent research [93], a multiobjective evolutionary approach was used for the simultaneous optimization of these two conflicting aspects during the identification of short interspersed repetitive elements in the DNA sequence of *Tripanosoma cruzi*: The method obtained all of the solutions identified by alternative single-objective approaches and discovered additional efficient trade-offs between the two objectives used [93]. MOO approaches to motif identification have also been explored in [94], [95], [96], with the aim of integrating several information sources [96], or to better specify the properties of the patterns sought [94], [95].

6.2 Structure Alignment

Due to the high evolutionary pressure on the structure of RNA and proteins which determines the function of these macromolecules, structural similarities are preserved to a much higher degree than are sequence similarities and may, consequently, still be identified in the absence of any apparent sequence similarities.

Analogously to the case in sequence alignment, structural alignment can be performed both on a global or a local level. Global sequence alignment serves to provide information on the evolutionary distance between macromolecular structures, usually to predict functional relationships or evolutionary relationships between molecules. In contrast, alignments on a local basis are more specifically aimed at the identification of shared functionally active regions, so-called pharmacophores. In applications involving local structure alignment, MOO approaches have been applied by a number of researchers [97], [98], [99], [100]. Here, the main two motivations for the use of multiobjective optimization were 1) to counterbalance the bias related to the length of the alignments compared [97], [99], [100] and 2) to simultaneously capture different aspects of solution quality [98].

An application related to pharmacophore identification is the search for median molecules that are representative for a given set of macromolecules [101]. The search for such median molecules, which will share properties with all of the target molecules, is an important technique in computer-aided molecular design [102]. Traditional approaches to the evolution of median molecules use a single-objective approach based on the sum of the distances to all target molecules or the distance to an average description molecule. As pointed out by Brown et al. [103], such an approach may be suboptimal as the objective value may be dominated by one of the target molecules and the resulting solution may be more representative of the corresponding molecule than of any of the other molecules. They therefore suggest a multiobjective evolutionary approach to the problem in which the distance to each target molecule is treated as a separate objective [103]. The advantages of the algorithm proposed were demonstrated in an application

involving the evolution of potential medians for sets of two target molecules.

7 Structure Prediction and Design

In this section, we continue to focus on optimization tasks related to the structure of macromolecules, in particular, the task of structure prediction and design. As mentioned previously, the functional properties of macromolecules derive from their three-dimensional shape (tertiary structure), which, in turn, is, predominantly, determined by their sequence of bases or amino acids (primary structure): Specifically, the native tertiary structure of a macromolecule is assumed to correspond to its lowest free-energy conformation. In theory, this direct relation between sequence and structure, which has been known for several decades, opens the door to *in silico* structure/function prediction (for a given sequence), as well as the design of new RNAs and proteins (for a given structure/function).

7.1 Protein Structure Prediction

In protein structure prediction, two fundamentally different approaches can be identified; those based on comparative modeling and those based on *de novo* modeling. Comparative modeling approaches predict structure based on that of homologous proteins: They are therefore only applicable if there are proteins with a high sequence similarity and known structure. In contrast to this, *de novo* modeling approaches can be applied to any protein sequence, but are, currently, less effective. The inherent difficulty of *de novo* protein structure prediction arises from two different issues: 1) the intricacy of formulating an energy function that realistically models the different local and global interactions contributing to protein folding and 2) the size of the space of possible conformations, which cannot be explored exhaustively. Progress in *de novo* protein structure prediction therefore crucially relies on progress both in the design of appropriate (i.e., more accurate) energy functions and the development of specialized efficient sampling methods.

Traditionally, empirical energy functions consist of a sum of the different energetic components contributing to the process of the folding of the macromolecule. Only recently have researchers aimed to identify the optimal weighting between these components using regression on a set of training data [104]. While this approach may lead to promising results, it is unclear whether such a fixed linear combination can provide optimal discrimination for all types of macromolecules and in all regions of the search space. Multiobjective optimization may therefore be a more principled approach and its use in protein structure prediction has recently been suggested by a number of authors [105], [106], [107].[3]

Schulze-Kremer [107] suggested the decomposition of the energy function into a nine-dimensional vector. Among others, the torsion energy, the van der Waals energy, the electrostatic energy, and a penalty energy term promoting compact folding patterns were taken into account and optimized using a multiobjective evolutionary algorithm. In [105], [106], a simpler two-objective formulation based on the CHARMM energy potential was proposed, where local and nonlocal interactions were treated as separate objectives. Promising results were obtained in a comparison to other algorithms across five different proteins [105].

A conceptually different approach has been suggested in [109]. Here, a weighted sum approach was used for the integration of different objectives in the prediction of protein structure from X-ray powder diffraction diagrams. Importantly, the objectives used related to fundamentally different types of information sources: One involved the minimization of the difference between the calculated and the measured diffraction patterns, while the other was based on the minimization of the potential energy of the system. In [109], the problem was optimized for a single weighting only (using single-objective simulated annealing), but, in subsequent work, a more rigorous exploration of different weights was proposed [110]. As discussed in [110], the attractiveness of this approach derives from the properties of the search landscapes created by the two individual objectives. The objective based on the diffraction is assumed to have a distinct global minimum, which can be unambiguously identified; unfortunately, the search for this minimum is hampered by the presence of multiple local minima and a small basin of attraction. In contrast, the objective based on free energy is assumed to have multiple local minima, which cannot be reliably differentiated, but have large basins of attraction; one of these minima can be expected to coincide with the global minimum in the first objective. The combination of both objectives may therefore serve as a way of facilitating the search problem by increasing the basin of attraction surrounding the optimal solution.

7.2 Directed Evolution

Directed evolution refers to the iterative production, evaluation, and selection of macromolecules in an *in vitro* environment. The aim of directed evolution is to evolve molecules with one or a number of desired functional properties by a process inspired by the process of selective breeding. As discussed previously, the function of a macromolecule is determined by its three-dimensional structure. While this can, in theory, be predicted from the sequence of a macromolecule, such rational design methods remain very limited in practice. Directed evolution (purportedly) avoids the need for efficient structure prediction methods through the evaluation of the performance of the macromolecules *in vitro* [111].

Recently, the success of techniques of directed evolution has been demonstrated by a number of researchers [112], [113], [114]. These studies have considered different types of macromolecules and objectives, but have been limited to a single objective. However, it may be argued that, in many applications, the aim of directed evolution may be better described as a multiobjective optimization problem, for example, if we aim to optimize the stability and the reaction rate of a given enzyme. A multiobjective evolutionary approach to the directed evolution of DNA sequences has recently been explored in [115]. The aim in this work was to obtain DNA sequences suitable for use in DNA computing, which results in a number of different objectives. Specifically,

3. Note that methods for protein and ligand docking use empirical energy functions very similar to those used in protein structure prediction and, thus, share the same motivation for a multiobjective approach [108].

the authors identified four different groups of objectives: those aimed at

1. the prevention of undesired reactions,
2. controlling the secondary structure,
3. controlling the chemical characteristics of the molecule, and
4. restricting the use of particular bases.

Using a multiobjective evolutionary algorithm, Shin et al. [115] implemented a six-objective version of the problem within a DNA sequence design system. The results obtained were compared to those obtained using traditional, single-objective formulations of the problem and, overall, the experiments indicated an increased robustness of the sequences generated using multiobjective optimization.

8 System Optimization and Experimental Design

In the previous section, we saw that multiobjective optimization can be useful in the design of improved macromolecules. It may also be used to investigate the degree of optimality of naturally occurring biochemical systems or to design optimal biochemical processes, and these ideas have been explored in a number of papers.

8.1 Study of the Optimality of Biochemical Systems

In [116], an experimentally derived kinetic model was optimized in order to study the trade-off between maximizing ethanol production in the yeast *Saccharomyces cerevisiae* while minimizing each of the internal metabolite concentrations. The optimization method employed was a linear programming approach, which was made publicly available [117].

The optimality of the heat shock response in cells was studied in [118]. Here, the term heat shock refers to the unfolding or misfolding of proteins, which is caused by sudden increases in temperature and is counteracted *inter alia* by an increased production of chaperones and proteases, which repair or degrade the damaged proteins. The costs related to this repair process and the costs related to the presence of damaged proteins can be seen as conflicting objectives, and El Samad et al. used a model of Differential Algebraic Equations and numerical optimization methods to identify the corresponding Pareto front. The results indicated that the heat response exercised by cells is close to optimal in the Pareto sense.

8.2 Optimization of Biochemical Processes

Optimal protocols for polymerization processes have been studied by a number of authors [13], [119], [120], [121]. Polymerization is the reaction process that joins single molecules into polymer chains and is of fundamental importance in the chemical and biochemical industries. Optimal polymerization is subject to a range of different conflicting objectives, including the polymerization degree and the reaction time, and it is thus naturally suited to a multiobjective approach. The problem has been approached using a number of different objectives and optimization algorithms.

Other examples of the use of multiobjective optimization for process optimization are applications related to the beer fermentation process [122], the citric acid fermentation of *Aspergillus niger* [14], the production of gluconic acid [123], as well as investigations regarding optimal liver function [124] and the production of oil in the yeast *Yarrowia lipolytica* [125].

8.3 Experimental Design

Combinatorial library design refers to the optimization of the collection of compounds to be used in a screening test for the identification of compounds that interact with a target enzyme or receptor. Despite the development of high-throughput screening tests, the proportion of potentially interesting compounds that can be subjected to the screening test in practice remains negligibly small. For this reason, a careful design of combinatorial libraries remains of fundamental importance in the field of drug design. It has been appreciated by a number of researchers that effective combinatorial drug design involves the optimization of a number of conflicting design criteria and may be best tackled using methods of multiobjective optimization [126], [127], [128].

The use of multiobjective approaches has also been investigated for a number of other design problems in bioinformatics, including the selection of single-nucleotide polymorphisms [129], oligonucleotide-design [130], [131], multiplex PCR assay design, and instrument optimization [132], [133].

9 Discussion and Outlook

In the previous sections, we have seen that multiobjective optimization has widespread applications in computational biology and bioinformatics. The performance gains and flexibility afforded by multiobjective optimization have been illustrated in a range of initial studies, but, in many of these problem domains, the full potential of multiobjective approaches in comparison to the current state-of-the art techniques remains to be explored.

The reasons underlying the use of multiobjective optimization differ widely across these application domains and we believe this aspect to be more interesting and revealing than a distinction between the specific techniques used. In Section 3, we put forward five different motivations for the use of multiobjective optimization and have referred to these different contexts along the way. Table 1 summarizes this classification of the different application areas. Note that several of the problems considered can potentially fall into more than one category, dependent on the specific viewpoint taken. In the table, we have only indicated those categorizations that correspond to views taken in the literature reviewed within the scope of this paper, but this categorization is clearly not final.

9.1 Visualization and Solution Identification

The large majority of MOPs we have identified in this review have been tackled by generating a whole Pareto front and by then applying (or hoping to apply) some form of decision-making process afterward to choose a single solution. This strategy defers decision making until "all of

TABLE 1
Categorization of the Main Applications Discussed

Problem	Standard MOO	Bias	Multiple data sources	Proxies	Multi--objectivization
ROC curves	—	—	—	√	—
Rule mining	—	—	—	√	—
Accuracy vs. complexity	—	—	—	√	—
Supervised feature selection	—	—	—	√	—
Ensemble learning	—	—	—	√	—
Clustering	—	√	√	√	—
Unsupervised feature selection	—	√	—	—	—
Association rule mining	—	—	—	√	—
Multidimensional scaling	—	—	√	√	—
Semi-supervision	—	—	√	—	—
Phylogenetic trees	—	—	√	—	—
Gene regulatory networks	—	—	√	√	—
Sequence alignment	—	√	√	√	—
Structure alignment	—	√	√	√	—
Structure prediction	—	—	√	√	√
Directed evolution	√	—	—	—	—
Biochemical systems	√	—	—	—	—
Biochemical processes	√	—	—	—	—
Experimental design	√	—	—	—	—

Some categories, such as "Multiobjectivization," are currently underrepresented, but we believe that further applications of this type will emerge as the use of multiobjective optimization propagates in the biological domain. The highest number of entries can currently be observed for the category "Proxy," which contains all those applications in which a "gold standard" (e.g., the best possible classifier over "future" data, the true structure of a protein, the true network of regulatory relationships, or the true evolutionary relationship between sequences/structures) exists, which we would like to reach, but do not have direct access to during the optimization process.

the information is in" (a good thing when little is known about the possible trade-offs), but the problem remains how to identify/select a single best solution. The really successful application of this mode of MOO thus calls for advanced methods for the visualization of the Pareto front and for the support of the decision maker in selecting solutions from it.

Evidently, straightforward visualizations of the Pareto front are only possible in two or three dimensions and a representation of the solutions obtained and the relationships between them becomes much more intricate for higher dimensions. To date, only a few methods for effective visualization have been introduced that can deal with the truly multidimensional case (one of the main examples is a parallel axis plot [134]) and visualization remains a major topic for future research.

Automatic identification of promising solutions from Pareto front approximations has been investigated in several recent works [135], [136], [137], [138], [139]. However, these papers have generally dealt with methods for steering/focusing the search toward the (potentially) more important areas without the need for additional preference information from the decision maker (usually by searching more strongly in regions of the Pareto front that have highest local curvature). An alternative approach is to first obtain the most complete Pareto front approximation set possible and then to, a posteriori, reduce this set to a single solution by some automated process, taking into account the whole Pareto front shape and other information. This approach has been investigated for several unsupervised classification tasks [62], [68], but remains to be explored in many other application domains.

More broadly, other approaches to support decision-making in MOO exist, too. Where expert knowledge on how to balance conflicting measures/goals is available, this can be extracted by using preference articulation techniques [18], [19], [140]. These have yet to be seen in biological applications, but we believe there is plenty of potential for future successes in this area.

10 Conclusion

This paper has outlined the wide applicability of multiobjective optimization in biological problem domains and has illustrated its potential with references to existing results from the literature, where available. Rather than differentiating between differences in the optimization techniques used, we have opted to emphasize differences in the reasons underlying the attractiveness of multiobjective approaches in different problem domains. We hope that this viewpoint will help to provide additional insight into the advantages afforded by multiobjective optimization with regard to the applications listed and/or additional problems encountered in the field.

Acknowledgments

J. Handl gratefully acknowledges the support of a doctoral scholarship from the German Academic Exchange Service and the Gottlieb Daimler and Karl Benz-Foundation, Germany. D.B. Kell would like to thank the BBSRC, EPSRC, NERC, and RSC for financial support. J. Knowles is supported by a David Phillips Fellowship from the Biotechnology and Biological Sciences Research Council (BBSRC), United Kingdom.

References

[1] J. Cohen, "Bioinformatics—an Introduction for Computer Scientists," *ACM Computing Surveys,* vol. 36, no. 2, pp. 122-158, 2004.
[2] S. Mitra, "Computational Intelligence in Bioinformatics," *Trans. Rough Sets,* pp. 134-152, 2005.

[3] D.F. Jones, S.K. Mirrazavi, and M. Tamiz, "Multi-Objective Meta-Heuristics: An Overview of the Current State-of-the-Art," *European J. Operational Research,* vol. 137, no. 1, pp. 1-9, 2002.
[4] S. Ruzika and M.M. Wiecek, "Approximation Methods in Multiobjective Programming," *J. Optimization Theory and Applications,* vol. 126, no. 3, pp. 473-501, 2005.
[5] C.A. Coello Coello, D.A. Van Veldhuizen, and G.B. Lamont, *Evolutionary Algorithms for Solving Multi-Objective Problems.* Kluwer Academic, 2002.
[6] D.W. Corne, K. Deb, P.J. Fleming, and J.D. Knowles, "The Good of the Many Outweighs the Good of the One: Evolutionary Multi-Objective Optimization," *Connections: The Newsletter of the IEEE Neural Networks Soc.,* vol. 1, no. 1, pp. 9-13, 2003.
[7] K. Deb, *Multi-Objective Optimization Using Evolutionary Algorithms.* John Wiley and Sons Ltd., 2001.
[8] C.M. Fonseca and P.J. Fleming, "Genetic Algorithms for Multiobjective Optimization: Formulation, Discussion and Generalization," *Proc. Fifth Int'l Conf. Genetic Algorithms,* pp. 416-423, 1993.
[9] M. Ehrgott and X. Gandibleux, "A Survey and Annotated Bibliography of Multiobjective Combinatorial Optimization," *OR Spectrum,* vol. 22, no. 4, pp. 425-460, 2000.
[10] *Multi-Objective Machine Learning, Studies in Computational Intelligence,* Y. Jin, ed., vol. 16, Springer-Verlag, 2006.
[11] R.T. Marler and J.S. Arora, "Survey of Multi-Objective Optimization Methods for Engineering," *Structural and Multidisciplinary Optimization,* vol. 26, no. 6, pp. 369-395, 2004.
[12] K. Deb and A.R. Reddy, "Reliable Classification of Two-Class Cancer Data Using Evolutionary Algorithms," *Biosystems,* vol. 72, nos. 1-2, pp. 111-129, 2003.
[13] S. Curteanu, F. Leon, and D. Galea, "Alternatives for Multiobjective Optimization of a Polymerization Process," *J. Applied Polymer Science,* 2006.
[14] C. Mandal, R.D. Gudi, and G.K. Suraishkumar, "Multi-Objective Optimization in Aspergillus Niger Fermentation for Selective Product Enhancement," *Bioprocess and Biosystems Eng.,* vol. 28, pp. 149-164, 2005.
[15] E.P. Someren, L.F.A. Wessels, E. Backer, and M.J.T. Reinders, "Multi-Criterion Optimization for Genetic Network Modeling," *Signal Processing,* vol. 83, pp. 763-775, 2003.
[16] M. Ehrgott, *Multicriteria Optimization.* Springer-Verlag, 2000.
[17] R. Steuer, *Multiple Criteria Optimization, Theory, Computation and Applications.* John Wiley and Sons Ltd., 1986.
[18] J.S. Dyer, P.C. Fishburn, R.E. Steuer, J. Wallenius, and S. Zionts, "Multiple Criteria Decision Making, Multiattribute Utility Theory the Next Ten Years," *Management Science,* vol. 38, no. 5, pp. 645-654, 1992.
[19] V. Belton and T.J. Stewart, *Multiple Criteria Decision Analysis: An Integrated Approach.* Springer-Verlag, 2002.
[20] H. Wright, K. Brodlie, and T. David, "Navigating High-Dimensional Spaces to Support Design Steering," *Proc. 11th IEEE Visualization Conf.,* pp. 291-296, 2000.
[21] J.D. Knowles, R.A. Watson, and D.W. Corne, "Reducing Local Optima in Single-Objective Problems by Multi-Objectivization," *Proc. First Int'l Conf. Evolutionary Multi-Criterion Optimization,* pp. 269-283, 2001.
[22] M.T. Jensen, "Guiding Single-Objective Optimization Using Multi-Objective Methods," *Applications of Evolutionary Computation,* pp. 268-279, 2003.
[23] C.E. Metz, "Basic Principles of ROC Analysis," *Seminars in Nuclear Medicine,* vol. 8, no. 4, pp. 283-298, 1978.
[24] M.A. Kupinski and M.A. Anastasio, "Multiobjective Genetic Optimization of Diagnostic Classifiers with Implications for Generating Receiver Operating Characteristic Curves," *IEEE Trans. Medical Imaging,* vol. 18, no. 8, pp. 675-685, 1999.
[25] R.M. Everson and J.E. Fieldsend, "Multi-Class ROC Analysis from a Multi-Objective Optimisation Perspective," *Pattern Recognition Letters,* vol. 27, pp. 918-927, 2006.
[26] L.A. Soinov, M.A. Krestyaninova, and A. Brazma, "Towards Reconstruction of Gene Networks from Expression Data by Supervised Learning," *Genome Biology,* vol. 4, no. 1, 2003.
[27] R. Bayardo and R. Agrawal, "Mining the Most Interesting Rules," *Proc. Fifth ACM SIGKDD Int'l Conf. Knowledge Discovery and Data Mining,* pp. 145-154, 1999.
[28] B. de la Iglesia, M.S. Philpott, A.J. Bagnall, and V.J. Rayward-Smith, "Data Mining Using Multi-Objective Evolutionary Algorithms," *Proc. IEEE Congress on Evolutionary Computation,* pp. 1552-1559, 2003.
[29] B. de la Iglesia, A.P. Reynolds, and V.J. Rayward-Smith, "Developments on a Multi-Objective Metaheuristic (MOMH) Algorithm for Finding Interesting Sets of Classification Rules," *Proc. Third Int'l Conf. Evolutionary Multi-Criterion Optimization,* pp. 826-840, 2005.
[30] F. Coenen, P. Leng, and L. Zhang, "Threshold Tuning for Improved Classification Association Rule Mining," *Proc. Pacific-Asia Conf. Knowledge Discovery and Data Mining,* 2005.
[31] V.N. Vapnik, *Statistical Learning Theory.* Springer-Verlag, 1998.
[32] W.S. Sarle, "Stopped Training and Other Remedies for Overfitting," *Proc. 27th Symp. Interface of Computing Science and Statistics,* pp. 352-360, 1995, ftp://ftp.sas.com/pub/neural/inter95.ps.Z.
[33] F. Esposito, D. Malerba, and G. Semeraro, "A Comparative Analysis of Methods for Pruning Decision Trees," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 19, no. 5, pp. 476-491, May 1997.
[34] A. Freitas, "A Critical Review of Multi-Objective Optimization in Data-Mining: A Position Paper," *SIGKDD Explorations,* vol. 6, no. 2, pp. 77-86, 2004.
[35] H. Ishibuchi and T. Yamamoto, "Fuzzy Rule Selection by Multi-Objective Genetic Local Search Algorithms and Rule Evaluatiion Measures in Data Mining," *Fuzzy Sets and Systems,* vol. 141, no. 1, pp. 59-88, 2004.
[36] X. Llora and D.E. Goldberg, "Bounding the Effect of Noise in Multiobjective Learning Classifier Systems," *Evolutionary Computation,* vol. 11, no. 3, pp. 278-297, 2003.
[37] D. Kim, "Structural Risk Minimization on Decision Trees Using an Evolutionary Multiobjective Optimization," *Proc. Seventh European Conf. Genetic Programming,* pp. 338-348, 2004.
[38] C. Igel, "Multiobjective Model Selection for Support Vector Machines," *Proc. Third Int'l Conf. Evolutionary Multi-Criterion Optimization,* pp. 534-546, 2005.
[39] S. Bleuler, M. Brack, L. Thiele, and E. Zitzler, "Multiobjective Genetic Programming: Reducing Bloat Using SPEA2," *Proc. IEEE Congress Evolutionary Computation,* pp. 536-543, 2001.
[40] E.D. De Jong, R.A. Watson, and J.B. Pollack, "Reducing Bloat and Promoting Diversity Using Multi-Objective Methods," *Proc. Genetic and Evolutionary Computation Conf.,* pp. 11-18, 2001.
[41] O. Nicolotti, V.J. Gillet, P.J. Fleming, and D.V.S. Green, "Multiobjective Optimization in Quantitative StructureActivity Relationships: Deriving Accurate and Interpretable QSARs," *J. Medical Chemistry,* vol. 45, pp. 5069-5080, 2002.
[42] H.A. Abbass, "Speeding Up Backpropagation Using Multiobjective Evolutionary Algorithms," *Neural Computation,* vol. 15, pp. 2705-2726, 2003.
[43] Y. Jin, T. Okabe, and B. Sendhoff, "Neural Network Regularization and Ensembling Using Multi-Objective Evolutionary Algorithms," *Proc. IEEE Congress Evolutionary Computation,* pp. 1-8, 2004.
[44] S. Wiegand, C. Igel, and U. Handmann, "Evolutionary Multi-Objective Optimization of Neural Networks for Face Detection," *Int'l J. Computational Intelligence and Applications,* vol. 4, no. 3, pp. 237-253, 2004.
[45] G.G. Yen and H. Lu, "Hierarchical Rank-Density Genetic Algorithm for Radial Basis Function Neural Network Design," *Int'l J. Computational Intelligence and Applications,* vol. 3, no. 3, pp. 213-232, 2003.
[46] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference and Prediction.* Springer-Verlag, 2001.
[47] V. Kumar and P. Rockett, "Multiobjective Genetic Algorithm Partitioning for Hierarchical Learning of High-Dimensional Pattern Spaces: A Learning-Follows-Decomposition Strategy," *IEEE Trans. Neural Networks,* vol. 9, no. 5, pp. 822-830, 1998.
[48] J. Liu and H. Iba, "Selecting Informative Genes Using a Multiobjective Evolutionary Algorithm," *Proc. World Congress Computational Intelligence,* pp. 297-302, 2002.
[49] J. Liu and H. Iba, "Prediction of Tumor Outcome Based on Gene Expression Data," *Wuhan Univ. J. Natural Sciences,* vol. 9, no. 2, pp. 177-182, 2004.
[50] L.S. Oliveira and R. Sabourin, "A Methodology for Feature Selection Using Multiobjective Genetic Algorithms for Handwritten Digit String Recognition," *Int'l J. Pattern Recognition and Artificial Intelligence,* vol. 17, no. 6, pp. 903-929, 2003.

[51] S.Y.M. Shi, P.N. Suganthan, and K. Deb, "Multi-Class Protein Fold Recognition Using Multi-Objective Evolutionary Algorithms," *Proc. IEEE Symp. Computational Intelligence in Bioinformatics and Computational Biology,* pp. 61-66, 2004.

[52] D. Stekel, *Microarray Bioinformatics.* Cambridge Univ. Press, 2003.

[53] J. Handl, J. Knowles, and D.B. Kell, "Computational Cluster Validation in Post-Genomic Data Analysis" *Bioinformatics,* vol. 21, no. 15, pp. 3201-3212, 2005.

[54] A.O. Hero, G. Fleury, A.J. Mears, and A. Swaroop, "Multicriteria Gene Screening for Analysis of Differential Expression with DNA Microarrays," *J. Applied Signal Processing,* vol. 1, pp. 43-52, 2004.

[55] A. Hero and G. Fleury, "Pareto-Optimal Methods for Gene Analysis," *J. VLSI Signal Processing,* vol. 38, no. 3, pp. 259-275, 2004.

[56] E. Bauer and R. Kohavi, "An Empirical Comparison of Voting Classification Algorithms: Bagging, Boosting, and Variants," *Machine Learning,* vol. 36, nos. 1-2, pp. 105-139, 1999.

[57] H.A. Abbass, "Pareto Neuro-Evolution: Constructing Ensemble of Neural Networks Using Multi-Objective Optimization," *Proc. IEEE Conf. Evolutionary Computation,* pp. 2074-2080, 2003.

[58] A. Chandra and X. Yao, "DIVACE: Diverse and Accurate Ensemble Learning Algorithm," *Proc. Fifth Int'l Conf. Intelligent Data Eng. and Automated Learning,* pp. 619-625, 2004.

[59] A.K. Jain and R.C. Dubes, *Algorithms for Clustering Data.* Prentice Hall, 1988.

[60] A. Strehl and J. Ghosh, "Cluster Ensembles—A Knowledge Reuse Framework for Combining Multiple Partitions," *J. Machine Learning Research,* vol. 3, pp. 583-617, 2002.

[61] A. Topchy, A.K. Jain, and W. Punch, "Clustering Ensembles: Models of Consensus and Weak Partitions," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 27, no. 12, pp. 1866-1881, Dec. 2005.

[62] J. Handl and J. Knowles, "An Evolutionary Approach to Multiobjective Clustering," *IEEE Trans. Evolutionary Computation,* vol. 11, no. 1, pp. 56-76, 2007.

[63] Y. Liu, T. Oezyer, R. Alhajj, and K. Barker, "Integrating Multi-Objective Genetic Algorithm and Validity Analysis for Locating and Ranking Alternative Clustering," *Informatica,* vol. 29, pp. 33-40, 2005.

[64] M.B. Dale and P.T. Dale, "Classification with Multiple Dissimilarity Matrices," *Coenoses,* vol. 9, no. 1, pp. 1-13, 1994.

[65] A. Ferligoj and V. Batagelj, "Direct Multicriterion Clustering," *J. Classification,* vol. 9, pp. 43-61, 1992.

[66] S. Madeira and A. Oliveira, "Biclustering Algorithms for Biological Data Analysis: A Survey," *IEEE/ACM Trans. Computational Biology and Bioinformatics,* vol. 1, pp. 24-45, Jan.-Mar. 2004.

[67] J.G. Dy and C.E. Brodley, "Feature Selection for Unsupervised Learning," *J. Machine Learning Research,* vol. 5, no. 5, pp. 845-889, 2004.

[68] J. Handl and J. Knowles, "Feature Subset Selection in Unsupervised Learning via Multiobjective Optimization," *Int'l J. Computational Intelligence Research,* vol. 2, no. 3, pp. 217-238, 2006.

[69] Y. Kim, W.N. Street, and F. Menczer, "Evolutionary Model Selection in Unsupervised Learning," *Intelligent Data Analysis,* vol. 6, no. 6, pp. 531-556, 2002.

[70] M. Morita, R. Sabourin, F. Bortolozzi, and C.Y. Suen, "Unsupervised Feature Selection Using Multi-Objective Genetic Algorithms for Handwritten Word Recognition," *Proc. Seventh Int'l Conf. Document Analysis and Recognition,* pp. 666-671, 2003.

[71] A. Freitas, "On Rule Interestingness Measures," *Knowledge-Based Systems J.,* vol. 12, nos. 5-6, pp. 309-315, 1999.

[72] A. Ghosh and B. Nath, "Multi-Objective Rule Mining Using Genetic Algorithms," *Information Sciences,* vol. 163, pp. 123-133, 2004.

[73] M. Khabzaoui, C. Dhaenens, and E.-G. Talbi, "A Multicriteria Genetic Algorithm to Analyze Microarray Data," *Proc. IEEE Congress Evolutionary Computation,* vol. 2, pp. 1874-1881, 2004.

[74] T.F. Cox and M.A. A. Cox, *Multidimensional Scaling.* Chapman and Hall, 2001.

[75] M. Brusco, "Multiobjective Multidimensional Scaling," *Proc. DIMACS Workshop Algorithms for Multidimensional Scaling,* 2001, http://dimacs.rutgers.edu/Workshops/.

[76] J. Weston, C. Leslie, D. Zhou, A. Elisseeff, and W. Noble, "Semi-Supervised Protein Classification Using Cluster Kernels," *Bioinformatics,* vol. 21, no. 15, pp. 3241-3247, 2005.

[77] J. Handl and J. Knowles, "On Semi-Supervised Clustering via Multiobjective Optimization," *Proc. Genetic and Evolutionary Computation Conf.,* pp. 1465-1472, 2006.

[78] J. Handl and J. Knowles, "Semi-Supervised Feature Selection via Multiobjective Optimization," *Proc. Int'l Joint Conf. Neural Networks,* pp. 6351-6358, 2006.

[79] D.B. Kell, "Metabolomics, Modeling and Machine Learning in Systems Biology—Towards An Understanding of the Languages of Cells. The 2005 Theodor Buecher Lecture," *FEBS J.,* vol. 273, pp. 873-894, 2006.

[80] R.D.M. Page and E.C. Holmes, *Molecular Evolution: A Phylogenetic Approach.* Blackwell Science, 1998.

[81] C. Cotta and P. Moscato, "Inferring Phylogenetic Trees Using Evolutionary Algorithms," *Proc. Seventh Int'l Conf. Parallel Problem Solving from Nature,* pp. 720-729, 2002.

[82] D. Felsenstein, *Inferring Phylogenies.* Sinauer Assoc., Inc., 2004.

[83] M. Salemi and A.-M. VanDamme, *Handbook of Phylogenetic Methods.* Cambridge Univ. Press, 2003.

[84] A. de Queiroz, M.J. Donoghue, and J. Kim, "Separate versus Combined Analysis of Phylogenetic Evidence," *Ann. Rev. Ecology and Systematics,* vol. 26, pp. 657-681, 1995.

[85] L. Poladian and L.S. Jermiin, "Multi-Objective Evolutionary Algorithms and Phylogenetic Inference with Multiple Data Sets," *Soft Computing,* vol. 10, pp. 359-368, 2006.

[86] P. Koduru, S. Das, S. Welch, and J.L. Roe, "A Multi-Objective GA-Simplex Hybrid Approach for Gene Regulatory Network Models," *Proc. IEEE Congress Evolutionary Computation,* pp. 2084-2090, 2004.

[87] C. Spieth, F. Streichert, N. Speer, and A. Zell, "Multi-Objective Model Optimization for Inferring Gene Regulatory Networks," *Proc. Conf. Evolutionary Multi-Criterion Optimization,* pp. 607-620, 2005.

[88] J.E. Fieldsend and S. Singh, "Pareto Evolutionary Neural Networks," *IEEE Trans. Neural Networks,* vol. 16, no. 2, pp. 338-354, 2005.

[89] R. Schwarz, P. Musch, A. von Kamp, B. Engels, H. Schirmer, S. Schuster, and T. Dandekar, "YANA—A Software Tool for Analyzing Flux Modes, Gene-Expression and Enzyme Activities," *BMC Bioinformatics,* vol. 6, no. 135, 2005.

[90] D.W. Mount, *Bioinformatics: Sequence and Genome Analysis.* Cold Spring Harbor Laboratory Press, 2004.

[91] M.A. Roytberg, M.N. Semionenkov, and O.Y. Tabolina, "Pareto-Optimal Alignment of Biological Sequences," *Biofizika,* vol. 44, no. 4, pp. 581-594, 1999.

[92] M.A. Roytberg, M.N. Semionenkov, and O.Y. Tabolina, "Sequence Alignment without Gap Penalties," *Proc. Int'l Conf. Bioinformatics of Genome Regulation and Structure,* 1998, http://www.bionet.nsc.ru/bgrs/thesis/85/.

[93] I. Zwir, R.R. Zaliz, and E.H. Ruspini, "Automated Biological Sequence Description by Genetic Multiobjective Generalized Clustering," *Annals New York Academy of Sciences,* vol. 980, pp. 65-82, 2002.

[94] V. Cotik, R.R. Zaliz, and I. Zwir, "A Hybrid Promoter Analysis Methodology for Prokaryotic Genomes," *Fuzzy Sets and Systems,* vol. 152, pp. 83-102, 2005.

[95] C.G. Nevill-Manning, T.D. Wu, and D.L. Brutlag, "Highly Specific Protein Sequence Motifs for Genome Analysis," *Proc. Nat'l Academy of Science,* vol. 95, pp. 5865-5871, 1998.

[96] M. Rajapakse, B. Schmidt, and V. Brusic, "Multi-Objective Evolutionary Algorithm for Discovering Peptide Binding Motifs," *Proc. Fourth European Workshop Evolutionary Computation and Machine Learning in Bioinformatics,* pp. 149-158, 2006.

[97] D. Baum, "Multiple Semi-Flexible 3D Superposition of Drug-Sized Molecules," *Proc. First Int'l Symp. Computational Life Sciences,* pp. 198-207, 2005.

[98] S.J. Cottrell, V.J. Gillet, R. Taylor, and D.J. Wilton, "Generation of Multiple Pharmacophore Hypotheses Using Multiobjective Optimisation Techniques," *J. Computer-Aided Molecular Design,* vol. 18, pp. 665-682, 2004.

[99] S. Handschuh, M. Wagener, and J. Gasteiger, "Superposition of Three-Dimensional Chemical Structures Allowing for Conformational Flexibility by a Hybrid Method," *J. Chemical Information and Computer Sciences,* vol. 38, pp. 220-232, 1998.

[100] Y. Jia, T.G. Dewey, I.N. Shindyalov, and P.E. Bourne, "A New Scoring Function and Associated Statistical Significance for Structure Alignment by CE," *J. Computational Biology,* vol. 11, no. 5, pp. 787-799, 2004.

[101] X. Jiang, A. Muenger, and H. Bunke, "On Median Graphs: Properties, Algorithms, and Applications," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 23, pp. 1144-1151, 2001.

[102] G. Schneider and U. Fechner, "Computer-Based de Novo Design of Drug-Like Molecules," *Nature Rev. Drug Discovery,* vol. 4, no. 8, pp. 649-63, 2005.

[103] N. Brown, B. McKay, F. Gilardoni, and J. Gasteiger, "A Graph-Based Genetic Algorithm and Its Application to the Multiobjective Evolution of Median Molecules," *J. Chemical Information and Computer Sciences,* vol. 44, pp. 1079-1087, 2004.

[104] J. Tsai, R. Bonneau, A. Morozov, B. Kuhlman, C.A. Rohl, and D. Baker, "An Improved Protein Decoy Set for Testing Energy Functions for Protein Structure Prediction," *Proteins,* vol. 52, pp. 76-87, 2003.

[105] V. Cutello, G. Narzisi, and G. Nicosia, "A Multi-Objective Evolutionary Approach to the Protein Structure Prediction Problem," *J. Royal Soc. Interface,* vol. 3, no. 6, pp. 139-151, 2006.

[106] R.O. Day, J.B. Zydallis, G.B. Lamont, and R. Pachter, "Solving the Protein Structure Prediction Problem through a Multiobjective Genetic Algorithm," *Nanotechnology,* vol. 2, pp. 32-35, 2002.

[107] S. Schulze-Kremer, "Application of Evolutionary Computation to Protein Folding with Specialized Operators," *Evolutionary Computation in Bioinformatics,* pp. 163-191, 2003.

[108] B. Liu, S. Mardikian, R. Jackson, and D. Westhead, "Computational Molecular Modeling of Protein Structure and Function," technical report, Univ. of Sheffield, Sheffield, U.K., 2004.

[109] H. Putz, J.C. Schoen, and M. Jansen, "Combined Method for Ab Initio Structure Solution from Powder Diffraction Data," *J. Applied Crystallography,* vol. 32, pp. 864-870, 1999.

[110] O.J. Lanning, S. Habershon, K.D. Harris, R.L. Johnston, B.M. Kariuki, E. Tedesco, and G.W. Turner, "Definition of Guiding Function in Global Optimization: A Hybrid Approach Combining Energy and R-Factor in Structure Solution from Powder Diffraction Data," *Physics Letters,* vol. 317, pp. 296-303, 2000.

[111] F.H. Arnold, "When Blind Is Better: Protein Design by Evolution," *Nature Biotechnology,* vol. 16, pp. 617-618, 1998.

[112] M.M. Altamirano, J.M. Blackburn, C. Aguayo, and A.R. Fersht, "Directed Evolution of New Catalytic Activity Using the Alpha/Beta-Barrel Scaffold," *Nature,* vol. 403, pp. 617-622, 2000.

[113] C. Schmidt-Dannert, D. Umeno, and F.H. Arnold, "Molecular Breeding of Carotenoid Biosynthetic Pathways," *Nature Biotechnology,* vol. 18, pp. 750-753, 2000.

[114] W.P.C. Stemmer, "Rapid Evolution of a Protein in Vitro by DNA Shuffling," *Nature,* vol. 370, no. 6488, pp. 389-391, 1994.

[115] S.-Y. Shin, I.-H. Lee, D. Kim, and B.-T. Zhang, "Multiobjective Evolutionary Optimization of DNA Sequences for Reliable DNA Computing," *IEEE Trans. Evolutionary Computation,* vol. 9, no. 2, pp. 143-158, 2005.

[116] J. Vera, P. d. Atauri, M. Cascante, and N.V. Torres, "Multicriteria Optimization of Biochemical Systems by Linear Programming: Application to Production of Ethanol by Saccharomyces Cerevisiae," *Biotechnology and Bioeng.,* vol. 83, no. 3, pp. 335-343, 2003.

[117] J. Vera and N.V. Torres, "MetMAP: An Integrated Matlab Package for Analysis and Optimization of Metabolic Systems," *Silico Biology,* vol. 4, no. 2, pp. 97-108, 2004.

[118] H. El Samad, M. Khammash, C. Homescu, and L. Petzold, "Optimal Performance of the Heat-Shock Gene Regulatory Network," *Proc. 16th IFAC World Congress,* 2005.

[119] K. Deb, K. Mitra, R. Dewri, and S. Majumdar, "Towards a Better Understanding of the Epoxy Polymerization Process Using Multi-Objective Evolutionary Computation," *Chemical Eng. Science,* vol. 59, no. 20, pp. 4261-4277, 2004.

[120] S. Garg and S.K. Gupta, "Multiobjective Optimization of a Free Radical Bulk Polymerization Reactor Using Genetic Algorithm," *Macromolecular Theory and Simulations,* vol. 8, pp. 46-53, 1999.

[121] K. Mitra, S. Majumdar, and S. Raha, "Multiobjective Dynamic Optimization of Epoxy Polymerization Process," *Computer and Chemical Eng.,* vol. 28, no. 12, pp. 2583-2594, 2004.

[122] B. Andres-Toro, J.M. Giron-Sierra, P. Fernandez-Blanco, J.A. Lopez-Orozco, and E. Besada-Portas, "Multiobjective Optimization and Multiobjective Control of the Beer Fermentation Process with the Use of Evolutionary Algorithms," *J. Zhejiang Univ. SCIENCE,* vol. 5, no. 4, pp. 378-389, 2004.

[123] H. Halsall-Whitney, D. Taylor, and J. Thibault, "Multicriteria Optimization of Gluconic Acid Production Using Net Flow," *Bioprocess and Biosystems Eng.,* vol. 25, pp. 299-307, 2003.

[124] N.S. Sharma, M.G. Ierapetritou, and M.L. Yarmush, "Novel Quantitative Tools for Engineering Analysis of Hepatocyte Cultures in Bioartificial Liver Systems," *Biotechnology and Bioeng.,* vol. 92, no. 3, 2005.

[125] L. Muniglia, L.N. Kiss, C. Fonteix, and I. Marc, "Multicriteria Optimization of a Single-Cell Oil Production," *European J. Operational Research,* vol. 153, no. 2, pp. 360-369, 2003.

[126] D.K. Agrafiotis, "Multiobjective Optimization of Combinatorial Libraries," *IBM J. Research and Development,* vol. 45, nos. 3/4, pp. 545-566, 2001.

[127] V.J. Gillet, "Designing Combinatorial Libraries Optimized on Multiple Objectives," *Methods in Molecular Biology,* vol. 275, pp. 335-54, 2004.

[128] V.J. Gillet, P. Willett, P.J. Fleming, and D.V. Green, "Designing Focused Libraries Using MoSELECT," *J. Molecular Graphics and Modeling,* vol. 20, no. 6, pp. 491-498, 2002.

[129] R.M. Hubley, E. Zitzler, and J.C. Roach, "Evolutionary Algorithms for the Selection of Single Nucleotide Polymorphisms," *BMC Bioinformatics,* vol. 4, no. 30, 2003.

[130] I.-H. Lee, S. Kim, and B.T. Zhang, "Multi-Objective Evolutionary Probe Design Based on Thermodynamic Criteria for HPV Detection," *Proc. Eighth Pacific Rim Int'l Conf. Artificial Intelligence,* pp. 742-750, 2004.

[131] C. Schretter and M.C. Milinkovitch, "OLIGOFAKTORY: a Visual Tool for Interactive Oligonucleotide Design," *Bioinformatics,* vol. 22, no. 1, pp. 115-116, 2006.

[132] M. Ehrgott and R. Johnston, "Optimisation of Beam Directions in Intensity Modulated Radiation Therapy Planning," *OR Spectrum,* vol. 25, no. 2, pp. 251-264, 2003.

[133] J. Knowles, "ParEGO: A Hybrid Algorithm with On-Line Landscape Approximation for Expensive Multiobjective Optimization Problems," *IEEE Trans. Evolutionary Computation,* vol. 10, no. 1, pp. 50-66, 2006.

[134] K.J. Shaw, A.L. Nortcliffe, M. Thompson, J. Love, P.J. Fleming, and C.M. Fonseca, "Assessing the Performance of Multiobjective Genetic Algorithms for Optimization of a Batch Process Scheduling Problem," *Proc. IEEE Congress on Evolutionary Computation,* pp. 37-45, 1999.

[135] J. Branke, K. Deb, H. Dierolf, and M. Osswald, "Finding Knees in Multi-Objective Optimization," *Proc. Eighth Int'l Conf. Parallel Problem Solving from Nature,* pp. 722-731, 2004.

[136] I. Das, "On Characterizing the 'Knee' of the Pareto Curve Based on Normal-Boundary Intersection," *Structural Optimization,* vol. 18, pp. 107-115, 1999.

[137] K. Deb, "Multi-Objective Evolutionary Algorithms: Introducing Bias among Pareto-Optimal Solutions," *Advances in Evolutionary Computing: Theory and Applications,* pp. 263-292, 2003.

[138] C.A. Mattson, A.A. Mullur, and A. Messac, "Smart Pareto Filter: Obtaining a Minimal Representation of Multiobjective Design Space," *Eng. Optimization,* vol. 36, pp. 721-740, 2004.

[139] G. Stehr, H. Graeb, and K. Antreich, "Performance Trade-Off Analysis of Analog Circuits by Normal-Boundary Intersection," *Proc. 40th Design Automation Conf.,* pp. 958-963, 2003.

[140] R.L. Keeney and H. Raiffa, *Decisions with Multiple Objectives: Preferences and Value Trade-Offs.* Cambridge Univ. Press, 1993.

[141] J. Handl, D.B. Kell, and J. Knowles, "Multiobjective Optimization in Bioinformatics and Computational Biology," Technical Report COMP-SYS-BIO-2006-04, Manchester Interdisciplinary Biocentre, Univ. of Manchester, U.K., Feb. 2006, http://dbk.ch.umist.ac.uk/handl/publications.html.

Julia Handl received the BSc (Hons) degree in computer science in 2001 from Monash University, Australia, the MSc degree in computer science in 2003 from the University of Erlangen-Nuremberg, Germany, and the PhD degree in 2006 from the University of Manchester, United Kingdom. She currently holds an MRC fellowship at the University of Manchester and works in the area of protein structure prediction.

Douglas B. Kell has been active for more than 10 years in the application of advanced computational techniques to the explanatory analysis of complex systems. He has published more than 350 scientific papers, holds six patents, and serves on a variety of bodies, including both the Council and the Strategy Board of the BBSRC and the NERC Environmental Genomics Committee, and has just completed a term on the RCUK Basic Technology Strategy Advisory Committee. He is a member of the EPSRC Peer Review College. He was founding director of Aber Instruments, which was set up to exploit his patented Biomass Monitor based on the use of noninvasive, radio-frequency dielectric spectroscopy. These instruments hold a world lead in this sector and Aber Instruments received the Queen's Award for Export Achievement in 1998. Since September 2002, he has held the EPSRC/RSC Research Chair in Bio-Analytical Science in the Department of Chemistry at UMIST, now the University of Manchester, with a brief to continue to develop and exploit novel analytical methods for complex postgenomic biosystems. He is the 2004 winner of the Royal Society of Chemistry Interdisciplinary Science Award, a 2005 Wolfson/Royal Society Merit Award, and the 2005 FEBS-IUBMB Theodor Buecher prize. He is also director of the Manchester Centre for Interdisciplinary Systems Biology.

Joshua Knowles received the BSc (Hons) degree in physics with mathematics in 1993, the MSc degree in information systems engineering (at the Distinction level) in 1996, and the PhD degree in 2002, all from the University of Reading, United Kingdom. From 2001-2003, he was a Marie Curie Fellow at IRIDIA, Free University of Brussels, and, since 2003, he has held a BBSRC David Phillips Fellowship at the University of Manchester, United Kingdom. His main research interest is heuristic, multiobjective optimization and its application to computational biology. He is currently serving on the UK EPSRC Peer Review College.