Quantitative Comparison of Approximate Solution Sets for Bi-criteria Optimization Problems*

W. Matthew Carlyle\(^\d\)
Operations Research Department, Navy Postgraduate School, Monterey, CA 93943, e-mail: mcarlyle@nps.navy.mil.

John W. Fowler, Esma S. Gel, and Bosun Kim
Arizona State University, Department of Industrial Engineering, P.O. Box 855906, Tempe, AZ 85287, e-mail: bosun.kim@asu.edu, john.fowler@asu.edu, esma.gel@asu.edu

ABSTRACT
We present the Integrated Preference Functional (IPF) for comparing the quality of proposed sets of near-pareto-optimal solutions to bi-criteria optimization problems. Evaluating the quality of such solution sets is one of the key issues in developing and comparing heuristics for multiple objective combinatorial optimization problems. The IPF is a set functional that, given a weight density function provided by a decision maker and a discrete set of solutions for a particular problem, assigns a numerical value to that solution set. This value can be used to compare the quality of different sets of solutions, and therefore provides a robust, quantitative approach for comparing different heuristic, a posteriori solution procedures for difficult multiple objective optimization problems. We provide specific examples of decision maker preference functions and illustrate the calculation of the resulting IPF for specific solution sets and a simple family of combined objectives.

Subject Areas: Evaluating the Quality of Approximate Solution Sets, Multiple Criteria Decision Making, and Multiple Objective Metaheuristics.

INTRODUCTION
Many real-world decision problems are appropriately modeled as multiple objective optimization problems. For example, in job scheduling, finding a sequence which minimizes the makespan, the sum of completion times, and the total weighted tardiness of the resulting schedule can be modeled as a multiple objective optimization problem. Problems of this sort (i.e., multiple objective combinatorial optimization) have become increasingly popular in the recent literature, although the single objective versions of these problems have been well studied (e.g., Pinedo, 1995, provides a review of many problems in scheduling). An illustration of the range of such problems and the types of criteria that might be used can be found in Steuer

\(^*\)This material is based upon work supported by the National Science Foundation under grant no. 0121815.
\(^\d\)Corresponding author.
Application areas of such problems are diverse: machine scheduling, production planning, network flows, containership loading, nurse scheduling, radiological worker allocation, and many more fall into the area of multiple objective optimization. In the management science literature, applications include vendor selection (e.g., Roodhooft & Konings, 1996), project scheduling (e.g., Viana & De Sousa, 2000), and investment decisions (e.g., Karsak & Kuzgunkaya, 2002) and could be extended in many ways to various workforce management issues (e.g., Karsak, 2001).

One of the research trends in multiple objective optimization, especially for problems that are NP-hard for each single objective, is to develop approximate solution algorithms using various metaheuristics such as evolutionary (genetic) algorithms, simulated annealing, tabu search, and so on. Coello Coello (1999) maintains a web site in which more than 100 journal papers and over 400 conference papers dealing with multiple objective heuristics are listed (see http://www.jeo.org/emo/). The majority of such heuristics are a posteriori solution approaches, which aim to generate a good approximation for all or a subset (e.g., supported solutions only) of the solutions in the efficient frontier (Coello Coello, 1999). The decision maker then implicitly compares the utility provided by each of the solutions and selects the most preferred single solution from the set. This approach is widely applied because it does not require the decision maker’s subjectivity during the solution procedure (Evans, 1984; Rosenthal, 1985).

One of the important (and unresolved) issues in a posteriori heuristics, as pointed out in Coello Coello (1999) and Ehrgott and Gandibleux (2000), is how to compare the performance of these various multiple objective metaheuristics fairly. The performance of an a posteriori heuristic can be evaluated in two general ways. One approach is to run competitive algorithms for the same amount of computational effort (CPU time or number of evaluations) and then compare the quality of the solutions. The other approach is to run each algorithm to its own stopping criterion and then compare both solution quality and computational efforts (Schaffer, 1985). In both cases, however, efficient methods to compare the quality of solution sets are required. Furthermore, as Hansen and Jaszkiewicz (1998) indicate, quantitative measures can be used to “tune” various parameters and determine stopping criteria based on solution quality in stochastic optimization methods. Finally, such measures can be used to trace the improvement of the solution quality of an approximation set.

Over 20 quantitative measures developed for this purpose can be classified into three main groups (see “Measures in the Literature” below). Among these, we are interested in the third group of measures, which use partial information on the decision maker’s value function. The basic concept of the measures in this group is that in an a posteriori solution approach, the decision maker eventually chooses the most preferred single solution among a set of nondominated solutions. To simulate this decision process, the form of the decision maker’s value function is assumed, although the specific parameters (i.e., the weight vector) are unknown. The uncertainty in the decision maker’s weight vector is represented via a density function. The expected utility (with respect to this preference density function) is calculated for each set of approximate solutions under consideration, and the approximation set with the best expected utility is determined as the winner. The reasoning behind this is that the final single solution chosen by the decision maker
is more likely to be found in the solution set with better expected utility. The measures in this group suggest Monte Carlo estimation to obtain the expected utility of approximation sets.

In this paper, we introduce a new measure called the integrated preference function (IPF), which is in some ways similar to the existing expected utility measures. However, we provide an analytical method of obtaining the exact expected utility when the parameterized combined objective function is a weighted additive function for bi-criteria problems. We also provide some intuitively desirable properties of IPF through numerical examples. Also, an illustration of our measure using the decision maker’s partial preference information (i.e., weight density function, weight range of objective, or priority of objectives) is provided.

In the “Measures in the Literature” section, measures that have previously been used in the literature are reviewed. The formalization of our approach, properties of our measure, the method to obtain the IPF under the value function mentioned above, and application of IPF using the decision maker’s partial preference information are provided in the section on IPF. In the “Numerical Examples” section, three numerical examples are presented. Finally, in the last section, we provide conclusions and discuss some topics that need further research.

MEASURES IN THE LITERATURE

In this section, we review the measures used for evaluating the quality of approximation sets generated by posteriori heuristics. Measures used for this purpose are summarized in Table 1. To our knowledge, there is no standardized framework (or a set of commonly applied measures) for fair comparison of heuristics, even though over 20 measures have been used in the literature.

Measures developed for this purpose can be classified into three groups. The first group aims to evaluate (or estimate) the quality of an approximate solution set based on the desirable attributes of approximation sets. A good approximation typically consists of a set of diverse solutions that are uniformly distributed along the efficient frontier, and which are also close to the efficient frontier. Also, a set with a higher number of solutions (i.e., greater cardinality) is preferred to a set with fewer solutions if both sets are generated with the same amount of computational effort. A diverse and uniformly distributed solution set with sufficient solution points can provide insights into the trade-offs between the objectives for the decision maker. Moreover, when the heuristics involve a stochastic nature (i.e., random numbers are used in its search mechanism), “less variance” among the output of replicates for the same problem instance is desired. Figure 1 illustrates the first three attributes for the minimization of two objectives. In each graph, solution set 1 is clearly better than solution set 2 in terms of the specified property.

However, there are difficulties associated with using measures in this first group. First, different researchers define desirable attributes differently (see Table 1). For example, the desirable number of solutions (cardinality) in a set is not clearly defined in the literature; too few solutions may be poor in providing the trade-off relation to the decision maker and too many solutions may overwhelm the decision maker while evaluating each solution to select the final solution. It is difficult to determine the proper cardinality of an approximation set without considering the way such a set is used in the decision-making process. Second, no
### Table 1: Measures that have previously appeared in the literature.

<table>
<thead>
<tr>
<th>Group</th>
<th>Subgroup</th>
<th>Measure Name</th>
<th>Desired</th>
<th>Quantity</th>
<th>EF/RS (^1)</th>
<th>Authors/year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measures intended to estimate each desirable attributes</td>
<td>Distance to the efficient frontier or reference set</td>
<td>Dist-1</td>
<td>Min</td>
<td>Yes</td>
<td>Czyzak &amp; Jaszkiewicz (1998)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Dist-2</td>
<td>Min</td>
<td>Yes</td>
<td>Czyzak &amp; Jaszkiewicz (1998)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Dist-z</td>
<td>Min</td>
<td>Yes</td>
<td>Viana &amp; De Sousa (2000)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Average Distance to Pareto optimal set</td>
<td>Min</td>
<td>Yes</td>
<td>Zitzler (1999)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Generational Distance</td>
<td>Min</td>
<td>Yes</td>
<td>Van Veldhuizen &amp; Lamont (2000)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Distance between ideal and nadir points</td>
<td>Max</td>
<td>No</td>
<td>Zitzler (1999)</td>
<td></td>
</tr>
<tr>
<td>Measures intended to estimate subjective solution quality based on Pareto dominance</td>
<td>Diversity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Uniformity</td>
<td>Spacing</td>
<td>-</td>
<td>No</td>
<td>Schott (1995)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td># Niche</td>
<td>Min</td>
<td>No</td>
<td>Zitzler (1999)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cardinality</td>
<td># APOS (^2)</td>
<td>Max</td>
<td>No</td>
<td>Schott (1995)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ratio of # APOS</td>
<td>Min</td>
<td>Yes</td>
<td>Van Veldhuizen &amp; Lamont (2000)</td>
<td></td>
</tr>
<tr>
<td>Measures Intended to estimate solution quality based on decision maker’s value function</td>
<td>Subjective</td>
<td>Visual Comparison</td>
<td>Implicit</td>
<td>No</td>
<td>Murata et al. (1996)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Area &amp; Length</td>
<td>Max</td>
<td>Yes</td>
<td>Zitzler (1999)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Size of Dominated Space</td>
<td>Max</td>
<td>No</td>
<td>Zitzler (1999)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Coverage difference</td>
<td>Max</td>
<td>No</td>
<td>Zitzler (1999)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Combined # APOS</td>
<td>Max</td>
<td>No</td>
<td>Schaffer (1985)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ratio of Combined # APOS</td>
<td>Max</td>
<td>No</td>
<td>Zitzler (1999)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Monte Carlo estimation (discrete weight)</td>
<td>Expected utility of the selection function</td>
<td>Min</td>
<td>Min</td>
<td>Esbensen &amp; Kuh (1996)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Expected Value (R1, R2, R3)</td>
<td>Min</td>
<td>No</td>
<td>Hansen &amp; Jaszkiewicz (1998)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Exact calculation (Continuous weight)</td>
<td>Maximum and average errors</td>
<td>Yes</td>
<td>Daniels (1992)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^1\)EF/RS: Is the efficient frontier or reference set required in comparing the quality of approximation sets? Some of the measures assume the efficient frontier is known. However, these measures are not always applicable since the efficient frontier is not known in many cases. In this case, one has to evaluate an approximate algorithm by comparing its results to those of other approximate algorithms.

\(^2\)APOS: approximate Pareto optimal solutions

measure in this group can be used alone; it is easy to construct, for each measure, an “optimal-looking” example for which the other measures are terrible. Hence, to evaluate the quality of approximation sets by using these desirable attributes, a multiple-objective problem should be solved. However, when the criteria are conflicting (i.e., when one set is better in the distance sense and the other set is better in the diversity sense), determining the winner is not easy, as can be seen in
Figure 1: Desirable properties of a set of approximate solutions (‘o’ represent set 1, a more desirable approximation, and ‘-’ represent set 2, an alternate approximation).

Figure 2: Examples of set dominance and set nondominance relations (‘o’ represents set $A^p$ and ‘-’ represents set $B^p$).

Fonseca and Flemming (1996). In Esbensen and Kuh (1996), a combined function of these criteria also did not provide the desired comparison results. Combining these attributes into a numerical score to rank competing heuristics may create further difficulties, especially in the integrated computerized framework.

The second group of measures aims to evaluate approximation sets based on the “set Pareto dominance” relation, which is defined below.

**Definition 1: Set Pareto Dominance**

Let $\{A^p\}$ and $\{B^p\}$ be approximation sets generated by competing heuristics, and $\{A^p\} \cup \{B^p\} = \{C\}$. Let $\{C^p\}$ denote the set of nondominated solutions from $\{C\}$. If $\{C^p\} \equiv \{A^p\}$ and $\{C^p\} \cap \{B^p\} \subset \{A^p\}$, then set $A^p$ is said to dominate set $B^p$ in set Pareto dominance sense and vice versa.

In Figure 2, graphs 1 through 4 are examples of the set Pareto dominance relation where set $A^p$ dominates set $B^p$. 
Definition 2: Nondominated Set Relation
Let \( \{ A^p \} \) and \( \{ B^p \} \) be approximation sets generated by competing heuristics, and \( \{ A^p \} \cup \{ B^p \} = \{ C \} \). Let \( \{ C^p \} \) denote the set of nondominated solutions from \( \{ C \} \). If \( \{ C^p \} \neq \{ A^p \} \) and \( \{ C^p \} \neq \{ B^p \} \), then set \( A^p \) and set \( B^p \) are said to be in a nondominated set relation.

In Figure 2, graphs 5 and 6 illustrate a nondominated set relation between set \( A^p \) and \( B^p \).

Clearly, if a set is dominated by another set or sets in the set Pareto dominance sense, there is no need to consider other attributes to compare the solution sets. However, in most cases, competing sets are in a nondominated set relation, which makes the comparison nontrivial.

There are five measures in this second group (see Table 1). Visual comparison is one of the most frequently employed methods in the literature, since all desirable attributes can be evaluated simultaneously, even though it involves human subjectivity. One important drawback of visual comparison, however, is that it can only be used properly in bi-criteria cases that can be depicted in a two-dimensional objective space. Furthermore, visual comparison might be practically impossible when a high number of computational tests are required to prove the superiority of approximate algorithms.

Among this second group, we can also list measures that use geometrical properties of solution sets. De et al. (1992) used area and length of approximately supported solutions. The authors assumed that both end points (individual optimal solutions) of the approximation set coincide with those of the efficient frontier. Again, these measures can only be used in bi-criteria cases and do not count on nonsupported solutions in an approximation set in evaluating solution quality. Zitzler and Thiele (1998) used the size of the dominated space of a set and the difference of the size of dominated spaces of two sets. This measure can evaluate the quality of approximations without knowledge of the efficient frontier. Also, this measure seems to evaluate the diversity and closeness simultaneously. However, in minimization problems, the assumption of the knowledge of the nadir point of the efficient frontier, which is often difficult to obtain, is needed to obtain the size of dominated space properly. Schaffer (1985) suggested a measure, which is referred to as the number of combined approximately Pareto-optimal solutions (#CAPOS). Zitzler (1999) used the ratio of #CAPOS of competing two sets. To obtain the #CAPOS of a set, all proposed approximate solutions are compared together in the Pareto dominance concept, and if any solution is dominated, it is discarded. Then, the number of approximate solutions found by each algorithm is determined. The cardinality and the ratio of #CAPOS are then used to indicate which algorithm is better in terms of solution quality. While the number of solutions generated is important, it certainly can be misleading.

The third group of measures evaluates the quality of solution sets based on their contribution to the decision-making process. The basic premise of these measures is that among a number of competing solution sets, the set that is more likely to contain the final single solution (to be selected by the decision maker) is determined as a better solution set. The methods assume that all solutions in a set
are provided to the decision maker sequentially or simultaneously and utilize partial information on the decision maker’s value function. For example, even though the weight for each objective may not be known, one can assume that the value function is a weighted additive function of the objectives. Daniels (1992), Esbensen and Kuh (1996), and Hansen and Jaszkiewicz (1998) have pioneered the use of measures in this group. However, the method suggested by Daniels (1992) is not applicable when the efficient frontier is not available. The methods suggested by Esbensen and Kuh (1996) and Hansen and Jaszkiewicz (1998), on the other hand, can be used when the efficient frontier is not known. However, rather than an exact computation scheme, a Monte Carlo estimation of expected utility (a discrete number of weight vectors generated randomly according to an assumed probability density function) was used in both methods, which clearly involve sampling errors (i.e., a weight vector for which a solution is optimal may not be sampled during the calculation of the expected value of an approximation set).

Three other types of measures, which are focused on a different aspect of a posteriori heuristics, can be found in the literature. Fonseca and Fleming (1996) proposed quantitative nonparametric interpretation of statistical performance of stochastic multiple objective optimizers. Laurnanns, Rudolf, and Schwefel (1999) suggested a measure (gain information) to trace the improvement of solution quality (convergence). Jaszkiewicz (2000) suggested a comparison method (effectiveness index) of computational effectiveness of a posteriori heuristics compared to single-objective metaheuristics used interactively.

INTEGRATED PREFERENCE FUNCTIONAL: IPF

In this section, we provide a formalization of our approach and a method to calculate the integrated preference functional (IPF) for a representative value function: the weighted additive value function for bi-criteria problems. We then discuss desirable properties of the IPF and finally show that IPF can be easily applied when partial information on the decision maker’s weights for different objectives is available. Throughout this paper we assume that all objectives have been transformed to be minimized.

Formalization of the IPF Measure

For multiple-objective optimization problems, the value (utility) function approach is frequently used to combine the various objective functions into one scalar function of the inputs. This combined objective can be represented as a parameterized family of functions $g(x; \alpha)$, where a given value of the parameter vector $\alpha$ in its domain $A$ represents a specific scalar-valued objective to be minimized. In the two-objective case $\alpha$ is a scalar between zero and one, for the case of a convex combination of objectives.

Given a finite set $X$ of proposed solutions (i.e., a set of nondominated solutions for the multiple objectives), there is at least one best solution (in this set) for the objective given by $g(x; \alpha)$ for any fixed value of $\alpha$. For a given $g$, define a function $x_g : A \rightarrow X$ that maps parameter values to a corresponding best solution in $X$. This function $x_g(\alpha)$ is clearly piecewise constant over $A$. The inverse function $x_g^{-1}(x)$,
Quantitative Comparison of Approximate Solution Sets

with \( x \in X \), defines a partition of parameter space \( A \) into the sets over which \( x_g \) is constant:

\[
A = \bigcup_{x \in X} x_g^{-1}(x) = \bigcup_{x \in X} A_x,
\]

for \( x_1 \neq x_2 \in X \), where \( A_{x_1} \) and \( A_{x_2} \) have at most one value in common in the two-objective case. In general, \( A_{x_1} \cap A_{x_2} \) is the boundary between the two regions for which \( x_1 \) and \( x_2 \) are optimal. This will, in all practical situations, be a set of measure zero, and will not impact the calculation of IPF.

Given a density function of the parameters \( h : A \to \mathbb{R}_+ \) such that \( \int_{\alpha \in A} h(\alpha) \, d\alpha = 1 \), we are interested in evaluating

\[
\text{IPF}(X) \equiv \int_{\alpha \in A} h(\alpha) g(x_g(\alpha); \alpha) \, d\alpha,
\]

the integrated preference functional that maps sets of solutions to the real numbers. Because \( x_g \) is piecewise constant, the integral in equation (1) can be decomposed into the portions of the domain \( x_g^{-1}(x) \) corresponding to each element \( x \in X \):

\[
\text{IPF}(X) = \int_{\alpha \in A} h(\alpha) g(x_g(\alpha); \alpha) \, d\alpha = \sum_{x \in X} \left[ \int_{\alpha \in x_g^{-1}(x)} h(\alpha) g(x; \alpha) \, d\alpha \right]
\]

Therefore, a given value of \( \alpha \) yields a particular objective function, for which there is at least one optimal solution in the set, \( x_g(\alpha) \). The density function \( h(\alpha) \) assigns different values to the weight vector \( \alpha \), and the IPF then provides a general measure of “optimality” of the set of solutions given, under the chosen weight density function.

The last form of the equations indicates that we only need to be able to evaluate integrals involving \( h(\alpha) \), and therefore the form of the objective functions themselves is irrelevant. Furthermore, the IPF measure proposed herein does not take any individual preference structures of the decision maker into account, and thus may be considered as generic. Of course, the main difficulties in all of this are computing the appropriate regions of \( A \) over which \( x_g \) is piecewise constant and calculating the integrals in equation (2). These difficulties depend on the type of the function \( g \), function \( h \), and the number of objectives considered.

**Weight Density Function \( h(\alpha) \)**

As mentioned before, the only requirement for IPF calculation is that \( h(\alpha) \) be an integrable function over \( \alpha \). This weight density function can be interpreted as the probability of decision maker’s preference for the weight of each objective. For example, if the decision maker does not have any specific preferences, then \( h(\alpha) \) can be modeled as a uniform density function as shown in Figure 3(a). In many cases, the decision maker may want to give more weight to well-compromised solutions (i.e., solutions located on the middle or elbow part of trade-off surface) than extreme solutions (i.e., solutions located near each individual optima). This can be modeled with a triangular density function as shown in Figure 3(b).
Figure 3: Two representative weight density functions for the decision maker’s (DM’s) preference for the weight of each objective.

Since the uniform density function $h(\alpha)$ can be expressed as

$$h(\alpha) = \begin{cases} 1, & 0 \leq \alpha \leq 1 \\ 0, & \text{otherwise} \end{cases}.$$  

IPF with uniform weight density function is obtained from the last form of the equation (2) as

$$\text{IPF}(X) = \sum_{x \in X} \left[ \int_{g \in x^{-1}(\alpha)} g(x; \alpha) \, d\alpha \right]$$

(3)

In the same way, since the triangular weight function $h(\alpha)$ can be expressed as

$$h(\alpha) = \begin{cases} 2\alpha/(1/2) = 4\alpha, & 0 \leq \alpha \leq 1/2 \\ 2(1 - \alpha)/(1/2) = 4(1 - \alpha), & 1/2 \leq \alpha \leq 1 \end{cases}.$$  

IPF with triangular weight density function is obtained from the last form of the equation (2) as

$$\text{IPF}(X) = \sum_{x \in X} \left[ \int_{0}^{1/2} 4\alpha g(x; \alpha) \, d\alpha + \int\limits_{1/2}^{1} 4(1 - \alpha) g(x; \alpha) \, d\alpha \right]$$

(4)

Properties of IPF of an Approximation Set

Some properties of IPF, which are intuitively desirable in evaluating the quality of solution sets, are provided below. As can be seen from equation (1), IPF considers only the best solution from an approximation set for the entire weight region.
**Theorem 1**
(IPF and set Pareto dominance relation): If set 1 is dominated by set 2 in the set Pareto dominance sense, then the IPF value of set 2 is less than or equal to the IPF value of set 1 for any positive weight function $h$.

**Proof**
Recall that the IPF value can be calculated by equation (2). When discrete numbers of weights are generated randomly from $h$, the following equation can be used to approximate the IPF value.

$$\text{IPF} \approx \frac{N}{N} \sum_{i=1}^{N} g(x; \alpha_i) / N,$$

where $N \to \infty$. Hence, the IPF is proportional to the minimum blended objective value (MBOV, or $g(x; \alpha)$) for each weight. If set 2 dominates set 1 in set Pareto dominance sense as in Definition 1, then MBOV of set 2 is always less than or equal to MBOV of set 1 in every weight (since we consider only a regular function $g$; i.e., if solution vectors $x_1 < x_2$, then $g(x_1; \alpha) < g(x_2; \alpha)$ for any value of $\alpha > 0$). Hence, the sum of MBOV of set 2 is always less than or equal to the sum of MBOV of set 1. Therefore, Theorem 1 holds. However, we note that the reverse does not always hold.

The equality can occur when an additive scalar function is used. If two sets have the same supported solutions and different nonsupported solutions, then the IPFs of both sets are the same. This is clearly a shortcoming in comparing approximate solution sets using a scalar function of weighted sum of linear objectives. However, this limitation can be removed by using nonconvex scalar functions. Theorem 1 implies the following corollary.

**Corollary 1**
(IPF of the efficient frontier and an approximation set): The IPF of the efficient frontier is always less than or equal to the IPF of an approximation set of the efficient frontier in a minimization problem.

**Proof**
The proof of this property is straightforward. The efficient frontier always dominates any approximation set in the set Pareto dominance sense. Hence, Corollary 1 holds from Theorem 1.

Corollary 1 implies that IPF can be used whether or not the efficient frontier is known. When the efficient frontier is known, the ratio of IPF of an approximation set over that of the efficient frontier or the IPF difference can be used to estimate how far an approximation set is from the efficient frontier, as indicated in Hansen and Jaszkiewicz (1998).

**Corollary 2**
(IPF and cardinality relation): Adding a new nondominated solution $x$ to the set $X$ can never increase the IPF value of the set $\{X + x\}$. Hence, IPF($X$) is monotonically nonincreasing over increasing sequences of solution sets.
Proof

Suppose that a newly added solution \( x \) is nondominated by any existing solution in set \( \{X\} \), and there is a weight \( \alpha_1 \) (or weight region) for which \( x \) is optimal for a given function \( g \). The IPF of the set \( \{X\} \) is proportional to \( \sum_{\alpha \in A} g(x; \alpha) \) as noted in Theorem 1. IPF of the set \( \{X + x\} \) is proportional to \( \sum_{\alpha \in A \setminus \alpha_1} g(x; \alpha) + g(x; \alpha_1) \). Clearly, \( g(x \in \{X + x\}; \alpha_1) \leq g(x \in X; \alpha = \alpha_1) \) by definition. Hence, the IPF of the set \( \{X + x\} \) is less than or equal to that of the set \( \{X\} \). If there is no such weight (or weight region) at which \( x \) is optimal for a given function \( g \), then the IPF of the set \( \{X\} \) and the set \( \{X + x\} \) are the same. Hence, Corollary 2 always holds.

By Corollary 2, IPF can be used to trace the solution quality improvement. As the measure (gain information) suggested by Laurnanns et al. (1999), let all individual optima \( \{X_{IO}\} \) be known in advance, and IPF \( \{X_{IO}\} \) denote the IPF of individual optima. By adding a nondominated solution \( x \) to the set \( \{X_{IO}\} \) sequentially, the IPF of the set \( \{X_{IO} + x\} \) will be monotonically nonincreasing by Corollary 2. By using this property of IPF, the convergence of an approximation can be estimated (i.e., the increment of IPF by adding solutions goes to 0). Furthermore, the net increment of IPF by adding a new solution to the set \( \{X_{IO}\} \) provides information about the solution quality improvement by adding a solution to an approximation set.

Scaling of Objectives

For applying IPF measure in real problems, proper scaling for each objective is needed. When the objectives considered are incommensurable (e.g., number of tardy jobs and total completion time), the blended objective value cannot be interpreted. Also, when the difference between the ranges of each objective value are so large that one objective value can be nullified by another objective value, proper scaling is necessary for blending multiple objectives to a reasonable scalar value. Schenkerman (1990) suggested that the proper minimum and maximum in scaling objective values is the nondominated minimum and ideal point of an approximation set respectively in a maximization problem. He also insisted that other minimums could prevent the decision maker from reaching a preferred decision. The same scaling method is employed in De et al. (1992) in comparing approximation sets with area and length measures for a minimization problem. As Gershon (1984) pointed out, scaling can be a measure of the importance of the objective, and this affects the weights considered.

Types of value functions

A variety of functions have been used as the decision maker’s value function, including linear additive functions (convex combination or weighted sum of the \( p \) objective function values), multiplicative functions, nonlinear additive functions (quadratic, square root, L4-norm), and multilinear functions (see Aksoy, Butler, and Minor, 1996, for a set of references for the use of each value function in comparing interactive methods). Lotfi, Yoon, and Zionts (1997) used linear, quadratic, and weighted Tchebycheff metrics to test their multiple objective linear programming (MOLP) algorithm. Among these, additive form and weighted Tchebycheff of scalar functions are most commonly used.
Quantitative Comparison of Approximate Solution Sets

IPF for a Weighted Additive Scalar Function

If we are using a convex combination of two objectives, $f_1$ and $f_2$, the function $g(x_\alpha(\alpha); \alpha)$ is the value of the blended objective $\alpha f_1(x_\alpha(\alpha)) + (1 - \alpha) f_2(x_\alpha(\alpha))$ evaluated at the optimal element of the efficient frontier (or the approximation set) for that particular objective. The decomposition in (2) then takes the form

$$IPF(X) = \int_{\alpha_0=0}^{\alpha_1} h(\alpha)g(x_\alpha(\alpha); \alpha) \, d\alpha + \int_{\alpha_1}^{\alpha_2} h(\alpha)g(x_\alpha(\alpha); \alpha) \, d\alpha + \cdots$$

$$+ \int_{\alpha_{k-1}}^{\alpha_k} h(\alpha)g(x_\alpha(\alpha); \alpha) \, d\alpha$$

$$= \int_{\alpha_0=0}^{\alpha_1} h(\alpha)[\alpha f_1(x_1) + (1 - \alpha) f_2(x_1)] \, d\alpha + \cdots$$

$$= \sum_{i=1}^{k} \left[ f_1(x_i) \int_{\alpha_{i-1}}^{\alpha_i} \alpha h(\alpha) \, d\alpha + f_2(x_i) \int_{\alpha_{i-1}}^{\alpha_i} (1 - \alpha) h(\alpha) \, d\alpha \right]$$

$$= \sum_{i=1}^{k} \left[ (f_1(x_i) - f_2(x_i)) \int_{\alpha_{i-1}}^{\alpha_i} \alpha h(\alpha) \, d\alpha + f_2(x_i) \int_{\alpha_{i-1}}^{\alpha_i} h(\alpha) \, d\alpha \right], \quad (5)$$

where $x_1$ is the value of $x_\alpha(\alpha)$ on the interval $(0, \alpha_1)$, $x_2$ is the value of $x_\alpha(\alpha)$ on the interval $(\alpha_1, \alpha_2)$, and so on. Note that here we assumed that there are $k$ intervals over each of which $x_\alpha$ is piecewise constant.

The following three steps illustrate the method of calculating IPF for the case of a bi-criteria optimization problem:

Step 1. Find all supported points and their adjacent supported points by the following algorithm:

- Sort $(f_1, f_2)$ points in increasing order of $f_2$ value;
- Let starting point $(f_1^0, f_2^0)$ be the first point in the sort list;
- Do until the last point in the sort list is selected as an adjacent point;
- Calculate the slope between the starting point $(f_1^0, f_2^0)$ and all the remaining points;
- Assign a point, which has the minimum slope to be the adjacent extreme point;
- Delete the nonsupported points between the starting point and the adjacent extreme point;
- Assign adjacent extreme point to be the new start point $(f_1^0, f_2^0)$;

Step 2. Find the optimal weight range for each supported point;

Step 3. Obtain IPF by equation $(5)$.

To perform step 2, assume $n$ supported solutions $x_k, k = 1, 2, \ldots, n$ are, obtained from step 1. Then, the optimal weight range can be obtained by the following
system of linear inequalities for each supported solution:

\[ \{ \alpha f_{1j} + (1 - \alpha) f_{2j} \} - \{ \alpha f_{1k} + (1 - \alpha) f_{2k} \} \leq 0, \quad j \neq k, j \in J, \]

where \( J \) is the set of adjacent supported points of \( x_k \).

In bi-criteria problems, every supported point has only two adjacent supported points except for the two tail points. Thus, two linear inequalities can be generated from two adjacent supported points. The two linear inequalities give the lower and upper bound of the optimal weight interval. The two tail points have one adjacent extreme point, which gives a bound of the optimal weight interval. The other bound is 0 or 1, since \( \alpha \) is assumed to be in \((0, 1)\).

The computational complexities of step 2 and step 3 in bi-criteria problems are \( O(m) \) each, and that of step 1 is \( O(n^2) \), where \( m \) is the number of supported solutions and \( n \) is the number of nondominated solutions. \( O(m) \) and \( O(n^2) \) are manageable polynomial-time functions; hence the computational effort to calculate IPF in the bi-criteria case is trivial.

**IPF with the Decision Maker’s Partial Preference Information**

Usually the decision maker does not know precisely the weight of each objective, but he or she may be able to specify some relations between weights. For instance, it may be easy for him or her to give an order on the importance of the objectives. If \( f_2(x) \) were more important than \( f_1(x) \), in effect, \( f_2(x) \succ f_1(x) \), then the weights would be \( 0 \leq \alpha_1 \leq \alpha_2 \). Sometimes, the decision maker can provide a weight interval (i.e., \( 0 \leq \alpha_1 \leq 0.4, 0.2 \leq \alpha_2 \leq 0.8 \)). These sorts of weight constraints can be easily incorporated with the IPF, since the weight constraints can be considered in obtaining optimal weight ranges and calculating integration.

As stated in Section 2, various types of weight density functions (i.e., uniform, triangular, and so on) can be incorporated with the IPF. Due to this characteristic of the IPF, it seems to be more suitable to compare guided multiple-objective metaheuristics (as in Branke, Kaussler, and Schmeck, 2000), than other measures. That is, a guided multiple-objective metaheuristics search for solutions in the region that the decision maker is interested in. If the decision maker’s preference information is available a priori, guided multiple-objective metaheuristics provide an a posteriori heuristic that utilizes this information to guide the search toward the area that is interesting for the decision maker a posteriori.

**NUMERICAL EXAMPLES**

Three numerical examples that give intuition for IPF are provided below:

**Example 1**

Comparing sets of approximate solutions using IPF with weighted additive value function and uniform weight density function.

In Figure 4, there are five sets of approximations. Each set (from set 1 to set 5) is denoted by a particular symbol (\( \Delta, \bullet, \square, \blacksquare, \) and \( \triangle \), respectively). The IPF value for each set calculated by equation \((3)\) is shown on the legend. By visual comparison, we see that set 1 clearly dominates all the other approximation
sets and set 5 is dominated by all four other sets. Sets 3 and 4 cross each other. Set 2 seems to be preferable to set 4 since most solutions in set 4 are dominated by solutions in set 2. However, the preference relation between set 3 and set 4 is not clear. Likewise, the preference relation between set 2 and set 3 is not clear.

From the IPF value for each set, the rank of each set can be determined easily. The best IPF value is 0.073 (set 1) and the worst IPF value is 0.411 (set 5). Set 2 has better IPF value (0.228) than those of set 3 (0.254) and set 4 (0.265). From these IPF values for five sets, the rank of each set is set 1 \( \succ \) set 2 \( \succ \) set 3 \( \succ \) set 4 \( \succ \) set 5. The IPF values of sets 1–4 and set 5 are an illustration of the first property, since set 5 is dominated by the other four sets and set 1 dominates the other four sets in set Pareto dominance sense. We can see that the IPF difference among sets in nondominated set relation is relatively small. The uniform weight density function implies that the decision maker has no specific region of interest; hence a set with lower IPF value is more likely to provide a better final solution on the average.

As shown in Example 1, it is clear that if an approximation set dominates another set in set Pareto dominance sense, the difference of IPF values for two sets is relatively large. However, if two sets do not dominate each other, it is not easy to decide which one is preferable for the decision maker even by visual comparison, since it depends on the decision maker’s preference structure.

**Example 2**
The IPF with weighted additive value function and triangular density function.

In Figure 5, two sets of approximations are compared using both a uniform weight density function (IPF-U) and a triangular weight density function (IPF-T).
When the two sets are compared using a uniform weight density function, the set represented by the circle symbol has a lower IPF value (0.288). On the other hand, if the two sets are compared with a triangular weight density function, the set marked by the heavy dash symbol has a lower IPF value (0.307). This indicates that when the decision maker is interested in the well-compromised solutions (which can be modeled as the triangular weight density function), a set that includes better solutions in the elbow area and worse solutions in extreme areas would have a lower IPF than a set that includes worse solutions in elbow area and better solutions in each extreme area.

**Example 3**
The IPF with partial preference information on weight.

In Figure 6, two sets of approximations are compared. When two sets are compared with IPF with weighted Tchebycheff metrics using a uniform weight density function, IPF values of both sets are the same (0.136). Assume that the decision maker’s preference that “$f_2(x)$ is more important than $f_1(x)$ (i.e., $f_1(x) \prec f_2(x)$)” is available. Then, this information can be converted into a weight interval of $0 \leq \alpha_1 \leq \alpha_2$. This constraint implies that $0 \leq \alpha_1 \leq 0.5$ (and $0.5 \leq \alpha_2 \leq 1$). When this information is used to obtain IPF values, the IPF of set 1 is 0.072 and the IPF of set 2 is 0.064, as shown in Figure 6. Hence, set 2 is determined as the winner by IPF measure. This implies that set 2 includes better solutions in the region that the decision maker is interested in than set 1.

**CONCLUSIONS AND FUTURE RESEARCH**

Many a posteriori heuristics have been developed for various multiple-objective optimization problems. Accordingly, over 20 different quantitative measures to evaluate the quality of approximation sets generated by such heuristics can be found in the literature. However, there is still no generally accepted measure or
Figure 6: Comparing two sets with preference information that objective-2 is more important than objective-1 ($f_1 < f_2$). “◦” represents set 1 and “-” represents set 2.

standard framework to compare solution quality of such heuristics. Considering the recent widespread use of multiobjective combinatorial optimization models for decision making, more research effort on the topic is warranted.

In this paper, we presented an exact measure (IPF) to evaluate the solution quality focusing on the decision-making process. We used a representative value function—weighted additive value function—to simulate the final single solution selection procedure by the decision maker among a set of alternative nondominated solutions.

Even though the suggested IPF measure cannot evaluate all desirable aspects of an approximation set explicitly, we have found that it provides fairly robust comparison results (see Carlyle, Kim, Fowler, & Gel, 2001; Kim, Gel, Carlyle, & Fowler, 2001) with respect to the desirable attributes. We also find that it is consistent with visual comparison, when such a comparison is easy. Furthermore, the IPF measure is easily applicable when partial information on the decision maker’s preference, such as priority of objectives or weight interval, is available. Due to this flexibility, IPF is more suitable in evaluating the quality of approximate solution sets, especially when the solution procedure is guided by the decision maker’s predetermined specific region of interests (relation between weights) as in Branke et al. (2000).

The IPF measure is intuitively more appealing to the decision maker when a posteriori heuristics are applied to solve real-world problems. Even though exact articulation of the decision maker’s value is generally very difficult, the type of value function, that is, additive value functions, seems to be fairly reasonable in many real-world problems. Usually, the decision maker experiences difficulty in the selection of a weight value for each objective. By providing the optimal weight range (or region) for each nondominated solution (which is a by-product of the IPF calculation), the decision maker can learn how the final optimal solution changes as his or her weight for each objective varies.

The most important assumption of IPF is that the decision maker’s preference information can be represented as a (convex or nonconvex) weighted additive value
function, which may or may not be true. We note, however, that our general framework is applicable to general forms of value functions, such as the Tchebycheff function. Our current research is on developing efficient algorithms to compute IPF for these general cases, as well as for problems with more than two objectives. One challenge that we are trying to overcome is that computational burden tends to increase significantly as the number of objectives increases, due to the computational inefficiency in both finding optimal weight regions and the high-dimensional integration. [Received: November 4, 2001. Accepted: December 23, 2002.]

REFERENCES


W. Matthew Carlyle is associate professor of operations research at the Naval Postgraduate School. He joined the faculty in 2002 after working as an assistant professor in the Department of Industrial Engineering at Arizona State University. He received his PhD in operations research from Stanford University in 1997 and his BS in information and computer science from Georgia Tech in 1992. His research interests include effective models and solution procedures for large combinatorial optimization problems. Applications of this research have included modeling and analysis of Navy combat logistics force size and structure, sensor mix and deployment for the Army’s objective force unit of action, workforce planning, underground mining operations, printed circuit-card assembly systems, and semiconductor manufacturing operations.

John W. Fowler received the PhD degree in industrial engineering from Texas A&M University. He is currently an associate professor of industrial engineering at Arizona State University (ASU) and is the director for the Factory Operations Research Center that is jointly funded by International SEMATECH and the Semiconductor Research Corporation. Prior to his current position, he was a senior member of technical staff in the Modeling, CAD, and Statistical Methods Division of SEMATECH and an adjunct assistant professor in the graduate program in operations research of the Mechanical Engineering Department at the University of Texas at Austin. His research interests include modeling, analysis, and control
of manufacturing (especially semiconductor) systems. He is the co-director of the Modeling and Analysis of Semiconductor Manufacturing Laboratory at ASU. The lab has had research contracts with NSF, SRC, International SEMATECH, Intel, Motorola, Infineon Technologies, ST Microelectronics, and Tefen, Ltd. Dr. Fowler is a member of American Society of Engineering Education, Institute of Industrial Engineers, Institute for Operations Research and Management Science, Production and Operations Management Society, and Society for Computer Simulation. He is an area editor for SIMULATION: Transactions of the Society for Modeling and Simulation International and an associate editor of IEEE Transactions on Electronics Packaging Manufacturing.

Esma S. Gel is currently assistant professor of industrial engineering at Arizona State University. Her research interests are multicriteria decision making, stochastic modeling and control of manufacturing systems, and workforce agility. She has received funding from the National Science Foundation, Semiconductor Research Corporation, and Infineon Technologies. She completed her PhD studies in 1999, at the Department of Industrial Engineering and Management Sciences of Northwestern University, where she also received her MS degree in 1996. She earned her BS degree in industrial engineering from Orta Dogu Technical University, Ankara, Turkey. She is a member of INFORMS, IIE, and ASEE.

Bosun Kim is a PhD candidate in the Department of industrial Engineering at Arizona State University. His research interests are in multicriteria decision making, heuristic algorithms, and production planning and scheduling, with applications to semiconductor manufacturing.